

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2005, 05:48:35 ; Search time 370 Seconds
(without alignments)
6337.256 Million cell updates/sec

Title: US-10-089-543-2
Perfect score: 1433
Sequence: 1 actataggcagcggtgtc.....tgagagaaatccttcacatc 1433

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113.8	7.9	31491	3	US-09-360-186-1
2	113.8	7.9	31491	4	US-09-864-680A-1
3	101.8	7.1	612	4	US-09-902-540-1357
4	90.2	6.3	59519	4	US-09-949-016-13504
5	88.8	6.2	55886	4	US-09-949-016-15129
6	88.4	6.2	19124	2	US-08-487-826B-13
7	88.4	6.2	205044	4	US-09-949-016-15851
8	88.4	6.2	205044	4	US-09-949-016-15852
9	88.4	6.2	205044	4	US-09-949-016-15853
10	88.4	6.2	223471	4	US-09-949-016-12387
11	88.4	6.2	223471	4	US-09-949-016-12724
12	88.4	6.2	223471	4	US-09-949-016-12725
13	87.4	6.1	61178	4	US-09-949-016-17369
14	85.8	6.0	187169	4	US-09-949-016-12776
15	85.8	6.0	191569	4	US-09-949-016-15940
16	84.4	5.9	1039	4	US-09-902-540-1280
17	84.2	5.9	263693	4	US-09-949-016-12386
18	84.2	5.9	263694	4	US-09-949-016-16915
19	84	5.9	20674	3	US-09-641-638-651
20	84	5.9	20674	4	US-10-170-037-651
21	83.8	5.8	147382	4	US-09-949-016-14624
22	83.6	5.8	601	4	US-09-949-016-30530
23	83.6	5.8	601	4	US-09-949-016-30531
24	83.6	5.8	601	4	US-09-949-016-37149
25	83.6	5.8	601	4	US-09-949-016-37150
26	83.6	5.8	601	4	US-09-949-016-37163
27	83.6	5.8	601	4	US-09-949-016-37164

28	83.6	5.8	601	4	US-09-949-016-145867	Sequence 145867, A
29	83.6	5.8	601	4	US-09-949-016-145868	Sequence 145868, A
30	83.6	5.8	601	4	US-09-949-016-146135	Sequence 146135, A
31	83.6	5.8	601	4	US-09-949-016-146136	Sequence 146136, A
32	83.6	5.8	601	4	US-09-949-016-146403	Sequence 146403, A
33	83.6	5.8	601	4	US-09-949-016-146404	Sequence 146404, A
34	83.4	5.8	615	3	US-08-998-416-186	Sequence 186, App
35	83.4	5.8	837	3	US-08-998-416-288	Sequence 288, App
36	82.6	5.8	19438	4	US-09-949-016-12699	Sequence 12699, A
37	81.6	5.7	18773	4	US-09-949-016-14164	Sequence 14164, A
38	81.6	5.7	95255	4	US-09-949-016-17067	Sequence 17067, A
39	81	5.7	18773	4	US-09-949-016-14164	Sequence 14164, A
40	80.8	5.6	636	3	US-08-998-416-1137	Sequence 1137, App
41	80.6	5.6	658	3	US-08-998-416-595	Sequence 595, App
42	80.2	5.6	29717	4	US-09-949-016-16284	Sequence 16284, A
43	80.2	5.6	60376	4	US-09-949-016-12423	Sequence 12423, A
44	79.8	5.6	134987	4	US-09-949-016-15348	Sequence 15348, A
45	79.8	5.6	134987	4	US-09-949-016-15349	Sequence 15349, A

ALIGNMENTS

RESULT 1
US-09-360-186-1
; Sequence 1, Application US/09360186
; Patent No. 6262343
; GENERAL INFORMATION:
; APPLICANT: Staskiewicz, et al.
; TITLE OF INVENTION: B2 Resistance Gene
; FILE REFERENCE: 50687
; CURRENT APPLICATION NUMBER: US/09/360,186
; CURRENT FILING DATE: 1999-07-23
; EARLIER APPLICATION NUMBER: 60/093,957
; EARLIER FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 31491
; TYPE: DNA
; ORGANISM: Capsicum annuum
; US-09-360-186-1

Query Match	7.9%	Score 113.8;	DB 3;	Length 31491;
Best Local Similarity	59.5%	Pred. No. 2.9e-14;		
Matches	210;	Conservative	0;	Mismatches 142;
			Indels	1;
			Gaps	1;
QY	38	CTTGATATCATGATTTTCAGATTGTCATTAAGACTTCTATCTATCAGAACGCTGCAG	97	
DB	22086	CTCTAGTTCTTCTTCGAAACATCGGCATAGGACATCTAACGACTTTTGAGCAGTCTTGAC	22145	
QY	98	AGGATCCCAATTAGTCTAAATTTATCTTCAGTCTCGAAACCAACTCAGGACCCAAAC	157	
DB	22146	TGCGTCTCAATGATTTTCTTCTTCATCGCCTTATGAAACAGATTAGGCCATACAA	22204	
QY	158	CCGTGCTCACCAACTCAGTCTAATATAACAGATGATGACATTTATGACATATAGC	217	
DB	22205	CTGAGTCTACCCACTTCATACCATCTATCGGAGACCTATATCTCTCAATACAAAGC	22264	
QY	218	CTGTAAGTGCCATCTAGATGCGAGATTGGAACCTGTTATTGAGGGGAACCTCACTAA	277	
DB	22265	CTCAAAAGGAGCCATCTTGATGCTGCGATGTTATTTATTTGTAACGGAATTCACCCAG	22324	
QY	278	CGGTAAATAATCTCTCAACTTCTAGTAATAATCACAATGCTCAAAATCGTATCTCTC	337	
DB	22325	TGGCAGGTGATCTACCCCACTACTCTTTGAAATCAATTCAGTCCTCCCTAACATATCTTC	22384	
QY	338	TAGTATATGATCACTCTTCTCAAAATTCACCATCGGTCTGAGGATGGAATCCAG	390	
DB	22385	GAGGCTGTAATGGTAGCCTCAGCTTGTCTCCATCGTCTGAGGGTGGAAACCTG	22437	

RESULT 2

US-09-864-680A-1
; Sequence 1, Application US/09864680A
; Patent No. 6762285
; GENERAL INFORMATION:
; APPLICANT: Staskiewicz, Brian J
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Tai, Thomas H
; TITLE OF INVENTION: B2 RESISTANCE GENE
; FILE REFERENCE: 42250/234021 (5830-4A)
; CURRENT APPLICATION NUMBER: US/09/864,680A
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/360,186
; PRIOR FILING DATE: 1999-07-23
; PRIOR APPLICATION NUMBER: US 60/093,957
; PRIOR FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 31491
; TYPE: DNA
; ORGANISM: Capsicum annuum
US-09-864-680A-1

Query Match 7.9%; Score 113.8; DB 4; Length 31491;
Best Local Similarity 59.5%; Pred. No. 2.9e-14;
Matches 210; Conservative 0; Mismatches 142; Indels 1; Gaps 1;
QY 38 CTGTGATCTCATGATTTTCAGATTGCGATTAAGACTTCTATCTATCAGAGACGCCCTGCAG 97
Db 22086 CTCTAGTCTCTCTCCGAAACATCGGCATAGGACATCTAAGACTTTTGAGCAGCTTGAC 22145
QY 98 AGGATCCCAATAGTCTAAATATCTTCAGTCTCGGAAACCAACTCAGGACCCCAAAAC 157
Db 22146 TCG-TCCTATGATTTTCATCTTCCTCCATGATTAAGCAAGATTAGGCCCATACAA 22204
QY 158 CCGTCGCTACCCCACTCAGCTTAATATAACAGATGACACTTATGACCATATAGAGC 217
Db 22205 CTGAGTCTACCCCACTCAGCTTAATATAACAGATGACACTTATGACCATATAGAGC 22264
QY 218 CTGATAGGTCCTATGATGCGAGATTCGAACTGTTATGAGCGAACTCAACTAA 277
Db 22265 CTCAAAAGGAGCCTCTGTATGCGATGCTGATGATTAATGTAAGCGAATTCACCCAG 22324
QY 278 CGGTAAATAATCCTCTCAACTACCTAGTAAATAATCACATAGTCCCAATCGTATCCTC 337
Db 22325 TGGCAGGTGATCTACCCCACTACCTTTGAAATCAATACCATGCCCTAAACATATCTC 22384
QY 338 TAGTATATGATCACCCTCTCTCAAAATGACCATCGGTCTGAGGATGGATCGAG 390
Db 22385 GAGGGTCTGAATGGTACGCTCAGCTTGCTCCATCGCTGAGGGTGGAAACCTG 22437

RESULT 3
US-09-902-540-1357/c
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,893
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:

; NAME/KEY: unsure
; LOCATION: (1)..(612)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357
Query Match 7.1%; Score 101.8; DB 4; Length 612;
Best Local Similarity 48.1%; Pred. No. 3.7e-12;
Matches 286; Conservative 0; Mismatches 309; Indels 0; Gaps 0;
QY 502 TTTTATTATTTTACCTCGCTAGATTGCTAAATAGTATGCAATTTATCTCAATTCATTAA 561
Db 611 TTTTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 552
QY 562 TTTTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 631
Db 551 TTTTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 492
QY 622 AGAATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 681
Db 491 TTTTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 432
QY 682 TAAARAAGAATTTCAATTCGTTTCTTAAATTTAGTTTAAATCTATATACTAATAATAA 741
Db 431 TTTTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 372
QY 742 AATTTCTGATCGGATTAGTGTGTCAGAAAGTCAAGTCACATGAATTTGTGGAGAAAAA 801
Db 371 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 312
QY 802 ATAAAAATTTAAACACATTTTCGATTAATTTATTTATTTATTTATTTATTTATTTT 861
Db 311 TTTTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 252
QY 862 TTATTTAAATGTTGCTCAATAATTTTAAATTTAAATTTTAAATTTTAAATTTTAAATTT 921
Db 251 TTTTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 192
QY 922 ATCAATTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 981
Db 191 TTTTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 132
QY 982 CACCCTCCATTAATGATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1041
Db 131 TTTTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 72
QY 1042 TTAATCATTAACCCAATTTTGAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1096
Db 71 TTTTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 17

RESULT 4
US-09-949-016-13504/c
; Sequence 13504, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13504
; LENGTH: 59519
; TYPE: DNA
; ORGANISM: Human

US-09-949-016-13504

```
Query Match          6.3%; Score 90.2; DB 4; Length 59519;
Best Local Similarity 48.8%; Pred. No. 2.9e-09;
Matches 301; Conservative 0; Mismatches 313; Indels 3; Gaps 2;

Qy 501 ATTTTATTTTACCTCCTGCTAGATTTCGTAATACTATTTCGCAATTTCTCATTTCAATT 560
Db 48491 ATAATATATATTATGTATGTTATACATATTATGTATAATATATGTTATATATATTATA 48432

Qy 561 ATTTATTTAATATTTTATATATTTTGGATFAAAAATTCATACTTTTACCTTTTAAAA 620
Db 48431 CATATAATGTATAATATATATATATATATATATATATATATATATATATATATATAT 48372

Qy 621 AAGAAATTTAATTTTATATATTTTATATATTTTATAGATAAAAATTCATACTTTTCTTTT 680
Db 48371 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 48312

Qy 681 TTAATAAGAAATTTCAATTCGCTTTTCTTAAATTTAGTTTAAATTTCTATACATAATTATA 740
Db 48311 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 48254

Qy 741 AAATTCGATCGGAATGATGCTGCAAGTCAGATCAATGAAATTTTGGTGAGAAAA 800
Db 48253 TATTATGTATAATATATATATATATATATATATATATATATATATATATATATATAT 48194

Qy 801 AATAAATAATTAACACATTTTTCGATTAATTTATATATATATATATATATATATATATAT 860
Db 48193 ATTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 48134

Qy 861 TTTATTTAATGCTGTCATATATATTTTAAATTTTAAATTTTTCAGCACAAATTTACACTCT 920
Db 48133 TATTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 48074

Qy 921 CATCAATTAATTTAATCTTTATACATAATTAATAATTTGAGGACAAATTTTAAATC 980
Db 48073 ATAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 48014

Qy 981 TCACCTCCATTAATGCATATATTAATTTT-TGTTCCATCTCTTTTACCTCTAA 1039
Db 48013 TTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 47954

Qy 1040 CATTAATCATTAACCAATTTTGAATCTTTAAATTTCTTAATTTTCACTTAATTTGTGGC 1099
Db 47953 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 47894

Qy 1100 TCTGGTCCATCTGGAA 1116
Db 47893 AATGTATATATGTAA 47877
```

RESULT 5

```
US-09-949-016-15129/c
; Sequence 15129, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15129
; LENGTH: 59886
; TYPE: DNA
```

ORGANISM: Human

```
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(55886)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15129
```

Query Match 6.2%; Score 88.8; DB 4; Length 55886;

Best Local Similarity 51.2%; Pred. No. 5.6e-09;

Matches 257; Conservative 0; Mismatches 242; Indels 3; Gaps 2;

```
Qy 538 TATTGCAATTTATCTCATTTTCATTTATTTTAAATTTATATATTTTATATATTTTGCATATAAAT 597
Db 5291 TATTTTATATATATAAAATCTTATTTTATATATATTTTATATATATTTTATATATATATAA 5232

Qy 598 CTAATACCTTTACTTTTCTTTTAAAAAGAAATTTTAAATTTTATATTTTATATTTTATAGATAAA 657
Db 5231 ATCATATTTTATATATATAAATCTTATTTTATATATATTTTATATATTTTATATATATAA 5172

Qy 658 AATTTCTAATCTTTTACTTTTCTTTTAAAAAGAAATTTCAATTCGCTTTTCTTAAATTTAG 717
Db 5171 ATCTTATTTTATATATATTTTATATATAAATAATCATATTTTATATATATATTTTATATAT 5112

Qy 718 TTTTAAATCTATATAATATATAAATTCGATCGGATTTAGTGGTGTCAAGTCAAGT 777
Db 5111 TTTATATATATAGAAATCTTATTTTATATATATTTTATATATATTTTATATATATATAA 5054

Qy 778 CACATGAATTTTGTGGAGAAAAATAAATAATTAACACACATTTTTCGATTTAATTTTATAT 837
Db 5053 AATCTTATTTTATATATATTTTATATATAATAAATCTTATTTTATATATATATTTTATAT 4994

Qy 838 ATATATAATATAATAACACATTTTATTTTAAATGTTGTCATAATATTTTAAATTAATAA 897
Db 4993 ATATATAAATCTTATTTTATATATATTTTATATATAATAAATCTTATTTTATATATA 4934

Qy 898 TTTGAGCACAAATTAACCTCTCATCAT-TAAATTTAACTTATTTACCATTAATTAATAAT 956
Db 4933 TTTTATATATATAATAAATCTTATTTTATATATATTTTATATATATAATAAATCTTATTT 4874

Qy 957 TGTGAGGACAAATTTTAAATCTCACCTCCATTAATGATATATTTTAAATTTTGTTC 1016
Db 4873 ATACAGCCATTTTCTAAATAAATAATACAGCCATTTTCTAAATAATATATTTTTCITT 4814

Qy 1017 GATACTTTCTTTTCACTCTCA 1038
Db 4813 TAAACTTTTTCCTTTTCCCA 4792
```

RESULT 6

```
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Welles, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 6.2%; Score 88.4; DB 2; Length 19124;
Best Local Similarity 49.8%; Pred. No. 5.2e-09;
Matches 289; Conservative 0; Mismatches 276; Indels 15; Gaps 2;

```
QY 429 TTAATATTTTAAATAATGATGTGACCAAGTGGTGGAGAGAGAGGTCTACCGATTGGT 488
Db 15951 TTTTATTTAAATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15892
QY 489 CAAGTGGCACCAATTTTATTTTACCTCGCTAGATTCGTAATACTATTGCAATTTA 548
Db 15891 TTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 15832
QY 549 TCTCATTTCAATTTATTTAAATTTATTTATTTATTTATTTGGATAAAATCTCAATCTTA 608
Db 15831 TTAATTTTATTTTAAATAAATTTTATTTTATTTTATTTATTTATTTATTTTATTTT 15776
QY 609 CTTTATTTTAAAGAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 668
Db 15775 ACATTTTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 15716
QY 669 TTTACATTTTATTTTAAAGAATTTCAATTTGCGTTTTTCTTAATTTAGTTTAAATCTTA 728
Db 15715 TTTTATTTTGTTTTATGATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15656
QY 729 TACTAATTAATAAAATTTCTGATCGGATAGTGGGTGTCAAAGTCAAGTCACATGAATTT 788
Db 15655 TTGTTTTTATTTTATATAATCATTTTATTTTATTTATATAAAATTTTATTTTATTTT 15596
QY 789 TGTGGAGAAAATAAATAAATAAACAATTTTCGATTAAATTTATTTATATATATATAATAA 848
Db 15595 TTGATAATCTTTTTCATTTTATTTCTATCAAAATTTATATTTTATTTATTTATTTATTA 15536
QY 849 TATAAACAATTTTATTTAATGTTGCAATAATTTTATTTTATTTTAAATTTTACGACAA 908
Db 15535 TTTTATAAATAATTTTCTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15486
QY 909 CAATTACACTCTCATCAATTAATAATTAATCTTATTAACATAATTAATAATTTGGAGGCAAT 968
Db 15485 AAATTTTATTTATTTATTTCTTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTT 15427
QY 969 TATTTTATTTATCTCACCTCCATTAATGTCATATTTAAT 1008
Db 15426 TCTTTTATTAATAATAATACATATATAATAAATAATATATAT 15387
```

RESULT 7
US-09-949-016-15851/c
; Sequence 15851, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15851
LENGTH: 205044
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(205044)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15851

Query Match 6.2%; Score 88.4; DB 4; Length 205044;
Best Local Similarity 49.1%; Pred. No. 9.2e-09;
Matches 288; Conservative 0; Mismatches 296; Indels 2; Gaps 2;

```
QY 427 AATATATTTTAAATAATGATGTGACCAAGTGGTGGAGAGAGAGGTCTACCGATTG 486
Db 201233 AACAAATTTTATTTGTAATCTGACTTATCTCCGAAAGGGATTTAGACAAACAC 201174
QY 487 GTCAAGTGGCACCAATTTTATTTTACCTCGCTAGATTCGTAATACTACTATTCGATT 546
Db 201173 TTCAATGTTTCATAGATAGGGTCAATTCCTCAITTTAAATATATTTCAAATATTAATAT 201114
QY 547 TATCTCATTTCAATTTATTTAAATTTATTTATTTATTTATTTGGATAAAATTTCTAAATCT 606
Db 201113 TATATTTAAATATATAAAATATATTTTATATATTTTATATATTTTAAATATATATTTTA 201054
QY 607 TACTTTTATTTAAAGAATTTTATTTTATTTTATTTTATTTTATTTAGATAAAATTTCTAAT 666
Db 201053 TATATTTATTTTAAATATATAATTTTATATATTTTATATATTTTATTTAAATATATATTT 200994
QY 667 ACTTTACTTTTATTTTAAAGA-ATTTCAAATTCGTTTTTCTTAATTTAGTTTAAAT 725
Db 200993 TATATATTTATTTTAAATATATATATTTTATATATTTTATATTTTAAATATATATAT 200934
QY 726 CTATACATTAATAAAATTCGATCGGATAGTGGTGTCAAGTCAAGTCACATGAA 785
Db 200933 TTTTATATTTTATTTTAAATATATATATTTTATATATTTTATATTTTAAATATATATA 200874
QY 786 TTTTGTGGAGAAAATAAATAAATAAACAATTTTCGATTAAATTTATTTATATATATAA 845
Db 200873 TTTTATATATTTATTTTAAAT-ATATATATTTTATATATTTTATATTTTAAATATAT 200815
QY 846 TAATATAAACAATTTTATTTAATGTTGCAATAATTTTATTTTAAATTTTAAATTTTCAGCA 905
Db 200814 ATATTTTATATATTTTATTTTAAATATATAAATATATATTTTATATATTTTATTTTAA 200755
QY 906 CAACAATTAACATCTCATCAATTAATTTAATCTTATTTACCATAATTTAAATTTGGAGGAC 965
Db 200754 AATATATAAATATATATTTTATATATTTTATATTTTAAATATATAAATATATTTTATAT 200695
QY 966 AATTTATTTTAAATCTCACCTCCATTTAATGTCATATTTTATTTTAAATTTT 1011
Db 200694 ATTATATTTTAAATATATAAATATATATTTTATATATTTTATTTTATTTT 200649
```

RESULT 8
US-09-949-016-15852/c
; Sequence 15852, Application US/09949016
; Patent No. 6812339

```

; Patent No. 6812339
;
; GENERAL INFORMATION:
;
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;
; FILE REFERENCE: CL001307
;
; CURRENT APPLICATION NUMBER: US/09/949,016
;
; CURRENT FILING DATE: 2000-04-14
;
; PRIOR APPLICATION NUMBER: 60/241,755
;
; PRIOR FILING DATE: 2000-10-20
;
; PRIOR APPLICATION NUMBER: 60/237,768
;
; PRIOR FILING DATE: 2000-10-03
;
; PRIOR APPLICATION NUMBER: 60/231,498
;
; PRIOR FILING DATE: 2000-09-08
;
; NUMBER OF SEQ ID NOS: 207012
;
; SOFTWARE: Fast-SEQ for Windows Version 4.0
;
; SEQ ID NO 15853
;
; LENGTH: 205044
;
; TYPE: DNA
;
; ORGANISM: Human
;
; FEATURE:
;
; NAME/KEY: misc_feature
;
; LOCATION: (1)...(205044)
;
; OTHER INFORMATION: n = A, T, C or G
;
; US-09-949-016-15853

```

	Query Match	6.2k;	Score 88.4;	DB 4;	Length 205044;
	Best Local Similarity	49.1k;	Pred. No. 9.2e-09;		
	Matches 288;	Conservative 0;	Mismatches 296;	Indels 2;	Gaps 2;
Qy	427	AAATTATTATTTTAAAAAATTGAATGACACAGTGGTTGGAGAGAGAGGTCTACCGATTG	486		
Db	201233	AACAAATTTTATTTGTTAAATCTGACTTTATTTCTCTGAAAGGGGATTTAGAGACAAACAC	201174		
Qy	487	GTCAAGTGGCCACCAATTTTTTATTTTAACTCTGCTAGATTTCGTAATACTATTTCGATT	546		
Db	201173	TTCATGTTTCATAGATGGGTCAATTCCTCATTTAAATATATATTCAAATATTAAATATT	201114		
Qy	547	TATCTCATTTTCATTATTATTAAATATTATTTATATTATTTGGATAAAAATTCATACTTT	606		
Db	201113	TATATTTAAAAATATATAAATATATTTTATATATTTATTTAAAAATATATATATTTTA	201054		
Qy	607	TACTTTTTTTTAAAAAGAAATTTATTTAAATTTTATATTATTTAGATAAAAATTCATAT	666		
Db	201053	TATATTTATATTTAAAAATATATATATTTTATATATTTATATTTAAAAATATATATTTT	200994		
Qy	667	ACTTTACTTTTTTTTTAAAAAGA-ATTTCAATTCGTTTTTTCTTTAAATTTAGTTTAAATT	725		
Db	200993	TATATATTTATTTTAAAAATATATATATTTTATATATTTTATATTTAAAAATATATATTT	200934		
Qy	726	CTATACAAATTATAAAAAATTCGTATCGGATTAGTGTGGTGTGCTCAAGTCACATGAA	785		
Db	200933	TTTATATATTTATATATTTAAAAATATATATATTTTATATATTTTATATTTAAAAATATATATA	200874		
Qy	786	TTTTTTGGAGAAAAAATAAAAAATTTAAACACATTTTTCGATTAAATTTATATATATATAA	845		
Db	200873	TTTTTATATTTTATATTTTAAAA-TATATATATTTTATATATTTATTTTAAAAATATAT	200815		
Qy	846	TAATATAAACACATTTTTTATTTTAAGTTGTGCAATAATATTTTTTAAATTTAAAAATTCAGCA	905		
Db	200814	ATATTTTTATATTTTATTTTAAAAATATATAAATATATTTTATATATTTTATATTTTAA	200755		
Qy	906	CAACAAATTTACACTCTCATCATTTAAATTTAACTTATTTACCATAATTTAAAAATGTGAGGAC	965		
Db	200754	AAATATATAAATATATTTTATATATTTTATATTTTAAAAATATATATAATATATTTTATAT	200695		
Qy	966	AAATTATTTTTTAACTCTCACCTCCCAATTAATGCATATTTTAAATTTT	1011		
Db	200694	ATTTTATTTTAAAAATATATAAAATATATTTTATATATTTTATATTTTAAATTTT	200649		

RESULT 9
US-09-949-016-15853/c
; Sequence 15853, Application US/09949016

```
Sequence 12387, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12387
LENGTH: 223471
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(223471)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12387

Query Match      6.2%; Score 88.4; DB 4; Length 223471;
Best Local Similarity 49.1%; Pred. No. 9.4e-09;
Matches 288; Conservative 0; Mismatches 296; Indels 2; Gaps 2;

QY 427 AATTATTATTTTAAAAAATTCATGTGACCGAGTGGTGGAGAGAGAGGCTACCGATTG 486
..DB 171660 AACAAATTTTATTTGTAATCTGACTTTATCTCCTGAAAGGGGATTAGAGACAACAC 171601

QY 487 GTCAAGTGGCCAAATTTTTTATTTTACCTCCCTGCTAGATTCGTAATAACTATTCGATT 546
DB 171600 TTCAATGTTTCATAGATAGGGGTCAAATTCCTCAATTTAAATATATTTCAAATATTAATATAT 171541

QY 547 TATCTCAATTCATTTTATTTTAAATTTATTTATTTATTTTGGATAAAAAATTCCTAATCT 606
DB 171540 TATATTTTAAATATATAAAATATATATTTTATATATTTTAAATATATATATATTTTAA 171481

QY 607 TACTTTTTTTTAAAAAGAAATTTTAAATTTTATTTTATTTTATTTAGATAAAAAATTCCTAAT 666
DB 171480 TATATTTTAAATATATAATATATATTTTATATATTTTATATTTTAAATATATATATTTT 171421

QY 667 ACTTTACTTTTTTTTAAAAAGA-AATTCAAATTCGGTTTTTCTTAAATTTAGTTTAAAT 725
DB 171420 TATATATTTTATATTTTAAAAATATATATATTTTATATATTTTATATTTTAAAAATATATAT 171361

QY 726 CTATACATAATATAAAATTCGATCGGATTCGATGGTGTGCTCAAGTCAAGTCACATGAA 785
DB 171360 TTTTATATTTTATTTTAAATATATATATTTTATATATTTTATATTTTAAATATATATATA 171301

QY 786 TTTTGTGGAGAAAAATAAAAAATTAACACATTTTTCGATTAATTTTATTTATATATATAA 845
DB 171300 TTTTATATATTTTATTTTAAAAA-TATATATATTTTATATATTTTATTTTAAAAATATAT 171242

QY 846 TAATATAAACACATTTTTTATTTAATGTGTCAATAATATTTTTTAAATTTAAATTTTCAGCA 905
DB 171241 ATATTTTATATATTTTATATTTTAAAAATATATAAAATATATTTTATATTTTATATTTAA 171182

QY 906 CAACAATTCACCTCATCAATTAATTTAATCTTTATTTACCATAATTAATAATTTGGAGGAC 965
DB 171181 AATATATATAAATATATTTTATATATTTTATATTTTAAATATATAAATAATATTTTATAT 171122

QY 966 AATTATTTTAAATCTCACCCCTCCATTAATGTCATATTTTAAATTTT 1011
DB 171121 ATTATATTTTAAATATATATAAATAATATTTTATATATTTTATATTTT 171076

US-09-949-016-12724/c
Sequence 12724, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12724
LENGTH: 223471
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(223471)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12724

Query Match      6.2%; Score 88.4; DB 4; Length 223471;
Best Local Similarity 49.1%; Pred. No. 9.4e-09;
Matches 288; Conservative 0; Mismatches 296; Indels 2; Gaps 2;

QY 427 AATTATTATTTTAAAAAATTCATGTGACCGAGTGGTGGAGAGAGAGGCTACCGATTG 486
DB 171660 AACAAATTTTATTTGTAATCTGACTTTATCTCCTGAAAGGGGATTAGAGACAACAC 171601

QY 487 GTCAAGTGGCCAAATTTTTTATTTTACCTCCCTGCTAGATTCGTAATAACTATTCGATT 546
DB 171600 TTCAATGTTTCATAGATAGGGGTCAAATTCCTCAATTTAAATATATTTCAAATATTAATATAT 171541

QY 547 TATCTCAATTCATTTTATTTTAAATTTATTTTATTTTATTTTGGATAAAAAATTCCTAATCT 606
DB 171540 TATATTTTAAATATATAAAATATATATTTTATATATTTTAAATATATATATATTTTAA 171481

QY 607 TACTTTTTTTTAAAAAGAAATTTTAAATTTTATTTTATTTTATTTAGATAAAAAATTCCTAAT 666
DB 171480 TATATTTTAAATATATAATATATATTTTATATATTTTATATTTTAAATATATATATTTT 171421

QY 667 ACTTTACTTTTTTTTAAAAAGA-AATTCAAATTCGGTTTTTCTTAAATTTAGTTTAAAT 725
DB 171420 TATATATTTTATATTTTAAAAATATATATATTTTATATATTTTATATTTTAAAAATATATAT 171361

QY 726 CTATACATAATATAAAATTCGATCGGATTCGATGGTGTGCTCAAGTCAAGTCACATGAA 785
DB 171360 TTTTATATTTTATTTTAAATATATATATTTTATATATTTTATATTTTAAATATATATATA 171301

QY 786 TTTTGTGGAGAAAAATAAAAAATTAACACATTTTTCGATTAATTTTATTTATATATATAA 845
DB 171300 TTTTATATATTTTATTTTAAAAA-TATATATATTTTATATATTTTATTTTAAAAATATAT 171242

QY 846 TAATATAAACACATTTTTTATTTAATGTGTCAATAATATTTTTTAAATTTAAATTTTCAGCA 905
DB 171241 ATATTTTATATATTTTATATTTTAAAAATATATAAAATATATTTTATATATTTTATATTTAA 171182

QY 906 CAACAATTCACCTCATCAATTAATTTAATCTTTATTTACCATAATTAATAATTTGGAGGAC 965
DB 171181 AATATATATAAATATATTTTATATATTTTATATTTTAAATATATAAATAATATTTTATAT 171122

QY 966 AATTATTTTAAATCTCACCCCTCCATTAATGTCATATTTTAAATTTT 1011
DB 171121 ATTATATTTTAAATATATATAAATAATATTTTATATATTTTATATTTT 171076

RESULT 11
```

RESULT 12

US-09-949-016-12725/c
; Sequence 12725, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12725
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12725

Query Match 6.2%; Score 88.4; DB 4; Length 223471;
Best Local Similarity 49.1%; Pred. No. 9.4e-09;
Matches 288; Conservative 0; Mismatches 296; Indels 2; Gaps 2;
Qy 427 AATTATTTATTTTAAAGAAATGATGACAGGTTGGAGAGAGGCTACCGATTG 486
Db 171660 AACAAATTTATTTGTAATCTGACTTATTTCTCTGAAAGGGATTTAGACAAACAC 171601
Qy 487 GTCAAGTGGACCAATTTTATTTTACCTCGCTAGATTCGTAATACATTTGCAAT 546
Db 171600 TTCATGTTTCAAGTAGAGGGTCAATTCCTCATTAATATATTTCAATATATATAT 171541
Qy 547 TATCTCAATTTCAATTTATTTAAATTTATTTATTTATTTTTCGATAAAATTTCTAAT 606
Db 171540 TATATTTAAATATATAATATATTTTATATATTTATTTTAAATATATATATTTTA 171481
Qy 607 TACTTTTTTTTAAAGAAATTTATTTAAATTTATTTATTTATTTAGATAAAATTTCTAAT 666
Db 171480 TATATTTATATTTAAATATATATATTTTATATATTTTATATTTTAAATATATATAT 171421
Qy 667 ACTTTACTTTTTTTTAAAGAA-ATTTCAATTCGGTTTTTCTTAATTTAGTTTAAAT 725
Db 171420 TATATATTTATTTTAAATATATATATATTTTTTATATATTTTAAATATATATAT 171361
Qy 726 CTATACTAATATATAAAATTTCTGATCGGATTAGTGTGTCAGAGTCAAGTCACATGAA 785
Db 171360 TTTATATATTTATTTTAAATATATATATTTTATATATTTATTTTAAATATATATA 171301
Qy 786 TTTTGTTCGAGAAATATAAAATTAACACATTTTTCGATTAATTTTATATATATATA 845
Db 171300 TTTTATATATTTTATTTTAAAT-TATATATATTTTATATATTTTATATTTTAAATATAT 171242
Qy 846 TAATATAACACATTTTATTTAAAGTGTGCAATATATTTTAAATTTAAATTTTCAGCA 905
Db 171241 ATATTTTATATTTTATTTTAAATATATATAATATATATTTTATATATTTTAAAT 171182
Qy 906 CAACAATTTACCTCATCAATTTAAATTTAATCTTATACCATAATTTAAATTTGAGGAC 965
Db 171181 AATATATATATATATTTTATATATTTTATATTTTAAATATATATAATATATTTTATAT 171122
Qy 966 AATTATTTTAAATCTCACCTCCCAATTAATGATATATTTTAAATTTT 1011
Db 171121 ATTTATATTTAAATATATATAAATATATTTTATATATTTTATATTTTATATTTT 171076

RESULT 13

US-09-949-016-17369
; Sequence 17369, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17369
; LENGTH: 61178
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17369

Query Match 6.1%; Score 87.4; DB 4; Length 61178;
Best Local Similarity 49.9%; Pred. No. 1.1e-08;
Matches 275; Conservative 0; Mismatches 271; Indels 5; Gaps 2;
Qy 546 TTATCTCATTTCAATTTATTTAAATTTATTTTATTTTATTTTGGATATAAAATTTCTAAT 605
Db 21250 TAATCTTATCTTATATAATCTTATATTTATTTATATATGCTCTTATATAATATATAT 21309
Qy 606 TTACTTTTTTTTAAAGAAATTTATTTAAATTTATTTTATTTTATTTAGATAAAATTTCTAA 665
Db 21310 ATAATCTTATCTTATATAATATAATCTTATTTATATAATAATAATCTTATTTATATAATA 21369
Qy 666 TACTTTTACTTTTTTTTAAAGAAATTTCAATTCGTTTTTCTTAATTTAGTTTAAAT 725
Db 21370 TCTTATCTTATATAATATAATATAATCTTATTTATATAATAATAATCTTATTTATATAAT 21429
Qy 726 CTATACTAATTTATAAAATTTCTGATCGGATTAGTGTGTCAGAGTCAAGTCACATGAA 785
Db 21430 ATAATCTTATATAATAATAATCTTATTTTATATAATAATCTTATTTATATAATAAT 21488
Qy 786 TTTTGTTCGAGAAATATAAAATTAACACATTTTTCGATTAATTTTATATATATAATA 845
Db 21489 ATAATCTTATCTTATATAATAATAATCTTATCTTATATAATAATAATAATCTTATCTTA 21548
Qy 846 TAATATAACACATTTTTTATTTTAAATTTGTCATAATAATTTTTTAAATTTAAATTTTCAGCA 905
Db 21549 TTATATAATAATAATCTTATCTTATTTATATAATAATAATAATAATAATAATCTTAACTTAT 21608
Qy 906 CAACAATTTACACTCTCATCAATTTAAATTTAATCTTTAT---TACCATTAATTTAAATTTGGA 961
Db 21609 TATATAATAATAATCTTATATAATAATCTTATCTTATATAATAATAATAATAATAATAAT 21668
Qy 962 GGCAATTTATTTTAAATCTCACCTCCATTAATGATATTTTAAATTTTGTTCGATAC 1021
Db 21669 ATATAATCTTATATTTTATATAATAATCTTATCTTATATAATAATGTTATTTATATAATA 21728
Qy 1022 TTTCTTATTTCACTCTCAACATTAATCAATTAACCAATTTTGAACCTGTTATATAATTTCTTAA 1081
Db 21729 TCTTATCTTATATAATAATGTTATTTATTTATATAATAATAATAATAATAATAATAATA 21788
Qy 1082 CTTATTTCACTA 1092
Db 21789 ATTATATATA 21799

RESULT 14
US-09-949-016-12776/c

Search completed: August 13, 2005, 13:15:41
Job time : 623 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2005, 11:00:09 ; Search time 8243 Seconds
(without alignments)
1127.927 Million cell updates/sec

Title: US-10-089-543-2

Perfect score: 1433

Sequence: 1 actataggcagcggtgtc.....tgagagaatccttcacatc 1433

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7305758 seqs, 3244068913 residues

Total number of hits satisfying chosen parameters: 14611516

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	137.8	9.6	4950	20	US-10-425-115-109152
C 2	136.4	9.5	4977	20	US-10-425-115-177410
C 3	133.6	9.3	4623	20	US-10-425-115-109144
C 4	132	9.2	4719	20	US-10-425-115-109143
C 5	131.2	9.2	3285	19	US-10-437-963-23515
C 6	131	9.1	4064	20	US-10-425-115-177422
C 7	129.4	9.0	2460	20	US-10-425-115-162914

C 8	129.4	9.0	4530	20	US-10-425-115-109139	Sequence 109139,
C 9	129.2	9.0	2373	19	US-10-437-963-78436	Sequence 78436, A
C 10	129.2	9.0	4641	19	US-10-437-963-78434	Sequence 78434, A
C 11	128.8	9.0	1155	20	US-10-425-115-177419	Sequence 177419,
C 12	128.8	9.0	4884	20	US-10-425-115-109131	Sequence 109131,
C 13	128.4	9.0	5433	20	US-10-425-115-177411	Sequence 177411,
C 14	128	8.9	58965	17	US-10-298-122-2	Sequence 2, Appli
C 15	126	8.8	4641	19	US-10-437-963-78438	Sequence 78438, A
C 16	124.8	8.7	4530	19	US-10-437-963-89508	Sequence 89508, A
C 17	123.6	8.6	1182	20	US-10-425-115-149480	Sequence 149480,
C 18	123.2	8.6	3969	19	US-10-437-963-77068	Sequence 77068, A
C 19	123.2	8.6	4303	19	US-10-437-963-89510	Sequence 89510, A
C 20	123.2	8.6	5292	19	US-10-437-963-89526	Sequence 89526, A
C 21	123.2	8.6	5430	19	US-10-437-963-89553	Sequence 89553, A
C 22	122.2	8.5	5162	19	US-10-437-963-89549	Sequence 89549, A
C 23	121.6	8.5	2781	19	US-10-425-115-111975	Sequence 111975, A
C 24	121.4	8.5	3057	20	US-10-437-963-77057	Sequence 77057, A
C 25	120	8.4	4638	19	US-10-437-963-89540	Sequence 89540, A
C 26	120	8.4	5058	19	US-10-437-963-89540	Sequence 89540, A
C 27	120	8.4	5208	19	US-10-437-963-89627	Sequence 89627, A
C 28	120	8.4	5547	19	US-10-437-963-89646	Sequence 89646, A
C 29	119.8	8.4	1326	19	US-10-437-963-34349	Sequence 34349, A
C 30	119.8	8.4	5199	19	US-10-437-963-89684	Sequence 89684, A
C 31	119.6	8.3	1056	20	US-10-425-115-109133	Sequence 109133, A
C 32	119.4	8.3	3936	19	US-10-437-963-85379	Sequence 85379, A
C 33	119	8.3	5043	19	US-10-437-963-14832	Sequence 14832, A
C 34	119	8.3	5151	19	US-10-437-963-23475	Sequence 23475, A
C 35	119	8.3	9603	19	US-10-437-963-54020	Sequence 54020, A
C 36	119	8.3	10549	19	US-10-437-963-89620	Sequence 89620, A
C 37	118.8	8.3	3849	20	US-10-425-115-149479	Sequence 149479, A
C 38	118.4	8.3	4092	19	US-10-437-963-89558	Sequence 89558, A
C 39	118.4	8.3	4149	19	US-10-437-963-89589	Sequence 89589, A
C 40	118.4	8.3	4611	19	US-10-437-963-77761	Sequence 77761, A
C 41	118.4	8.3	5281	19	US-10-437-963-85626	Sequence 85626, A
C 42	118.4	8.3	5448	19	US-10-437-963-89685	Sequence 89685, A
C 43	118.4	8.3	5556	19	US-10-437-963-89725	Sequence 89725, A
C 44	118.4	8.3	5571	19	US-10-437-963-89616	Sequence 89616, A
C 45	118.4	8.3	6101	19	US-10-437-963-29850	Sequence 29850, A

ALIGNMENTS

RESULT 1
US-10-425-115-109152/c
; Sequence 109152, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 109152
; LENGTH: 4950
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_31042C.1
; US-10-425-115-109152

Query Match 9.6%; Score 137.8; DB 20; Length 4950;
Best Local Similarity 62.2%; Pred. No. 5.2e-10;
Matches 217; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
Qy 58 GATTTCATAGACTTCTATCTATCAGAGCGCTCAGAGGATCCCAATTAGTCTAA 117
Db 4492 GGTCTGCATACTCTCTGACGAGCTTGGGCTTTCTTATATTAATTAATCTATGAA 4433

```
QY 118 AATTATCTTCAGTCTCGGAACCAACTCAGGACCCCAAAACCCCGTCGCTCACCACACTCAG 177
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4432 CTTTTCTTCACCTCTTCCACCATATCAGGTCTAAGAGATATCTTCTCCAGGTTCCAG 4373
QY 178 TCTAATATAACAGAGATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCATCTAGA 237
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4372 ACCAAATTAAAGGAGTTCGACACCGAGTCCATATAAAGCTTCAAAGGGTGCCATCTTGA 4313
QY 238 TGCAGATTGGAACCTGTTATTGTAGCGCAACTCAACTAAGCGTAAATAATCTCTCAAC 297
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4312 TACTTTCTGGTAACTGTTATTGTATGAAAACTCCGCTAGGGCAACATTCATCCATT 4253
QY 298 TACCTTAGTAATAATCACATAGCTCCAAATCGTATCTCTAGTATATGAATCACCTTCT 357
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4252 TTGGTGAGAAGTCCAGGACACATGCCCGCAGCATATCTTCAAGTATTTGGTTCACCTCT 4193
QY 358 CAAATTGACCATCGGTCTGAGGATGGAATGCAGACCGGTGCCACCGATT 406
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4192 CAGTTTGTCCACTGCTTTGAGGATGATAGGCCGAACCTGTGGAGCAGTTT 4144
```

RESULT 2

```
US-10-425-115-177410
; Sequence 177410, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 177410
; LENGTH: 4977
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_93387C.1
US-10-425-115-177410
```

```
Query Match 9.58; Score 136.4; DB 20; Length 4977;
Best Local Similarity 62.7%; Pred. No. 8.2e-10;
Matches 212; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 60 TTTGTCATAAGACTTCTATCTATCAGAAGACGCCCTGCAGAGAGATCCCAAAATTAGTCTAAAA 119
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
461 TCTGCATAACTCTTTTCCGCTGCTTGAGCTTTCTTCATATATGGAATAATCTCTGAACC 520
QY 120 TTATCTTTCAGTCTCGGAACCAACTCAGGACCCCAAAACCCGCTCACCACACTCACTGTC 179
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
521 TTTTCTTTCACCTCTTCCACCATATCAGGTCTAAAGAAGTATCTTTCTCCAGGTTCCAGAC 580
QY 180 TAATATAACAGAGTATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCATCTAGATG 239
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
581 CAAATTTAACGGAGTTCAGACACCGACGTCATATAAAGCTTCAAAGGGTGCCATCTTGATA 640
QY 240 CCAGATTGGAACCTGTTATTGTAGCGCAACTCAACTAAGCGTAAAAAATCTCTCAACTA 299
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
641 CTTTCTTGTAACTGTTATTGTATGAAAATTCGCTAGGGGTAAACATTATCCCATTTT 700
QY 300 CTTTGTAAATAAATCAATAGTCCAAATCGTATCTCTAGTATATGAATCACCTTCTCA 359
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
701 GGTGAGAAATCCAGGACACATGCCCGCAGCATATCTTCTAGTATTTGGTTCACCTCTCA 760
QY 360 AATTGACCATCGGTCTGAGGATGGAATGCAGACCGGTG 397
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
761 GTCTGCCCACTGGTTTGAGGATGATAAGCCGAACCTGTG 798
```

```
RESULT 3
US-10-425-115-109144/c
; Sequence 109144, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 109144
; LENGTH: 4623
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_31035C.1
US-10-425-115-109144
```

```
Query Match 9.3%; Score 133.6; DB 20; Length 4623;
Best Local Similarity 62.1%; Pred. No. 2e-09;
Matches 211; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 58 GATTTCATAGACTTCTATCTATCAGAAGACGCCCTGCAGAGATCCCAAAATTAGTCTAA 117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4165 GGTCTGCATAACTCTTTTGTGAGCCTGAGCTTTCTCATATATATGAACAAATTTCTGAA 4106
QY 118 AATTATCTTCAGTCTCGGAACCAACTCAGGACCCCAAAACCCCGTCGCTCACCACACTCAG 177
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4105 CTTTCTTTCGACCTCTTTCACCATATCAGGCTTAAGAAATATCTTCTCCAGGTTCCAG 4046
QY 178 TCTAATATAACAGAGTATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCATCTAGA 237
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4045 ACCAATTAGCGGAGTTCGACACCGTCCATATAAAGCTTCAAAAGGTGCCATCTTGA 3986
QY 238 TGCAGATTGGAACCTGTTATTGTAGCGCAACTCAACTAAGCGTAAAAAATCTCTCAAC 297
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3985 TACTTTCTTGATAAAGTATGTAAGAAACTCCGCTAGGGCAACATTCATCCCATTT 3926
QY 298 TACCTTAGTAATAAATCACATAGCTCCAAATCGTATCTCTAGTATATGAATCACCTTCT 357
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3925 TTGCTGAGAAATCCAGGACACATGCCCGCAGCATATCTTCAAGTATTTGGTTCACCTCT 3866
QY 358 CAAATTGACCATCGGTCTGAGGATGGAATGCAGACCGGTG 397
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3865 CAGTCTGTCCACTGGTTTGAGGATGATAAGCCGAACCTGTG 3826
```

RESULT 4

```
US-10-425-115-109143/c
; Sequence 109143, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 109143
; LENGTH: 4719
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
```

OTHER INFORMATION: Clone ID: MRT4577_31034C.1
US-10-425-115-109143

Query Match 9.2%; Score 132; DB 20; Length 4719;
Best Local Similarity 61.8%; Pred. No. 3.4e-09;
Matches 210; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Qy 58 GATTTCGATGAACTTCTATCTATCAGAGACGCTGCGAGAGATCCCAATTTAGTCTAA 117
Db 4261 GGTCTGATTAACCTTTTGTGAGTCTGAGCTTCTTCATATTATGATTAATTTCTGTAA 4202

Qy 118 AATTATCTTCAGTCTCGAAACCAACTCAGGACCCAAACCGTGGCTCAACCACATCTCAG 177
Db 4201 CCTTTCTTCGACCTCTTTCCACCATATCAGGCGCTTAAGAAAGTATCTTTCTCCAGGTTCA 4142

Qy 178 TCTAATAATACAGAGTATGACACTTATGACCATATAGAGCCTCGTAAGTGCCATCTAGA 237
Db 4141 ACCAATTTAGCGAGTTTCGACACTGTCTCCATATAAAGCTTCAAAAGGTGCCATCTTGA 4082

Qy 238 TGCAGATTGGAACCTGTTATTTAGGCGAACTCAACTAAACGGTAAATAATCTCTCAAC 297
Db 4081 TACTTCTTGATTAAGTCTGTTATGTAAGAAATTCGCTAAGGCGAATTCATCCATT 4022

Qy 298 TACCTTAGTAAATACATAGCTCCAAATCGTATCTCTAGTATATGAATCACTTCT 357
Db 4021 TTTGCGAGAGTCCAGAACACATGCGCGAGCATATCTCGAGTATCTGATTCACCTCT 3962

Qy 358 CAATTCACCATCGTCTGAGGATGGAATGACGACCGGTG 397
Db 3961 CAGTCTGCTCACTGTTTGGAGTATGATAGCGCAACTGTG 3922

RESULT 5

US-10-437-963-23515/c
Sequence 23515, Application US/10437963
Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 23515
LENGTH: 3285

TYPE: DNA

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_28588C.1
US-10-437-963-23515

Query Match 9.2%; Score 131.2; DB 19; Length 3285;
Best Local Similarity 61.3%; Pred. No. 4e-09;
Matches 211; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

Qy 52 TTTTCAGATTGCAATAGACTTCTATCTATCAGAGACGCTGCGAGGATCCCAATTA 111
Db 1953 TTCTGACATCGCGTAGCTCTCTCTCTACTTTGAGCGCGAAGCATGTGTGCGGTATCT 1894

Qy 112 GTCTAAATATCTTTCAGTCTCGGAACCACTCAGGACCCAAACCGCTCGCTCACCA 171
Db 1893 CTGCTATCCGTTCTGTGTGTGTTGAACCCCAATCGGTCCTTAAATAGTCTCTCCCTTA 1834

Qy 172 ACTCAGTCTAATAACAGAGTATGACACTTATGACCATATAGAGCCTCGTAAGTGCCA 231

Db 1833 TGGATCCCAACACAAATGGGATTCGACATTTTCTCCATACAGTGGCTCATATGGTGCCA 1774
Qy 232 TCTAGATGCGAGATTGGAACCTGTTATTTGTAGGCGAACTCAACTAAACGGTAAATAATCTCT 291
Db 1773 TTCGAATGCTAGCATGTAACATTTCTTGTGAACAAATTTCTACGAGCGCAATGATCTCT 1714
Qy 292 CTCACATCTCTTAGTAAATTAATCACAATAGCTCCAAATCGTATCTCTTAGTATATGAATCA 351
Db 1713 CCCAACTGCGCTTTCCCATGAAATTTATACATGATCTCAACATATCTCTTAGAGTTTGAATGG 1654
Qy 352 CCTTCTCAATTCACCATCGTCTGAGGATGGAATGCAGACCGG 395
Db 1653 TCGCTCAGATTGACCATCAGTCTGAGGGTGAAGCCGCTACTG 1610

RESULT 6

US-10-425-115-177422
Sequence 177422, Application US/10425115
Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 177422
LENGTH: 4064

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: MRT4577_93398C.1
US-10-425-115-177422

Query Match 9.1%; Score 131; DB 20; Length 4064;
Best Local Similarity 61.1%; Pred. No. 4.5e-09;
Matches 212; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

Qy 60 TTTGCATAAGACTTCTATCTATCAGAGACGCTGCGAGGATCCCAATTTAGTCTTAAA 119
Db 461 TCTGCATAACTCTTTTGGCGTGTGAGCTTTCTCATATTATGGATAAATTTCTGAAAC 520

Qy 120 TTATCTTCAGTCTCGGAACCAACTCAGGACCCAAACCGCTCACCNACTCAGTC 179
Db 521 TTTTCTTCAACCTCTTTTACCATATCAGGTCTTAAAGAAAGTATCTTCTCCAGGTTCA 580

Qy 180 TAATAAACAAGATGATGACACTTATGACCATATAGAGCCTCGTAAGTGCCATCTAGATG 239
Db 581 CAATTTAGCGAGTCCGACACCGCTGTCATATAAAGCTTCAAGGTCGCACTTGTATA 640

Qy 240 CCAGATTGGAACCTGTTTATTTGTAGGGGAACCTCAACTAACCGTAAATAAATCTCTCAACTA 299
Db 641 CTTTCTTGTAGTATTTATTTATGAAAACCTCCGCTAAAGGTAACATTCATCCATTTT 700

Qy 300 CCTTAGTAATAATACATAGCTCCAAATCGTATCTCTTAGTATATGAATCACCCTTCTCA 359
Db 701 AGTGAGAAATCCAGGACACATGCGCGCAGCATATCTTCAAGTATTTGGTTCACCTCTCA 760

Qy 360 AATTGACCATCGCTGAGGATGGAATGACAGCCGCGCCACCGATT 406
Db 761 GTTTGTCCACTGTTTGGAGTATAGGCCGAACCTGTTGGAGCAGTTT 807

RESULT 7

US-10-425-115-162914/c
Sequence 162914, Application US/10425115
Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 162914
; LENGTH: 2460
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_80152C.1
US-10-425-115-162914

Query Match          9.0%; Score 129.4; DB 20; Length 2460;
Best Local Similarity 61.4%; Pred. No. 6.6e-09;
Matches 208; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 59 ATTTGCATAGACTTCTATCTATCAGAGAGCGCTGCAGAGGATCCCAAAATTAGTCTAAA 118
DB 2001 ATCTGGCGTAACCTCTTCTGAGGAGCTTGAGCTTTCTCAAAATTATGAATTATTCGTTGAAC 1942

QY 119 ATTATCTTCAGTCTCGGAACCACTCAGGACCCAAAACCCGCTCGCTCACCCCAACTCAGT 178
DB 1941 CTTCCTCTCTGCTCTCTTTACCATATCAGGCTGAGAAATATCTTTCACCAAGTTTCAGA 1882

QY 179 CTAATATAACAGAGTATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCATCTAGAT 238
DB 1881 CCAGTTTAAACGAGTGCAACATCGTCTGCATATATAAAGCTTCGAAGGTGCCATCTTAAT 1822

QY 239 GCAGATTGAAACTGTTATGTAGGCGAACTCAACTAAGCGTAAATAATCTCTCAACT 298
DB 1821 ACTTCTTGATAGCTATTATATATGAGAACTCCGCTAATGGTAAACAATCATCCCATTT 1762

QY 299 ACCTTAGTAATAATACATAGCTCCAAATCGTATCTCTAGTATATGAATCACTTCTC 358
DB 1761 CTGTGGGAATCTCAGAACACATGCTCGCAACATATCTTCGAGTATCTGATTACCTCTC 1702

QY 359 AAATTGACCATCGTCTGAGGATGGAATGAGACCGGTG 397
DB 1701 AGTCTGCCACTGCTTTGAGGATGATAGGCGCAACTATG 1663

RESULT 8
US-10-425-115-109139/c
; Sequence 109139, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 109139
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_31030C.1
US-10-425-115-109139

Query Match          9.0%; Score 129.4; DB 20; Length 4530;
Best Local Similarity 60.8%; Pred. No. 7.7e-09;
Matches 211; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
```

```
QY 60 TTTGCATAAGACTTCTATCTATCAGAGAGCGCTGCAGAGGATCCCAAAATTAGTCTAAA 119
DB 4070 TCTGCATAACTCTTTTGGCGTCTTGAGCTTTCTTCATATATGGAATAATTCTCTGAACC 4011

QY 120 TTATCTTCAGTCTCGGAACCACTCAGGACCCAAAACCCGCTCGCTCACCAACTCAGTC 179
DB 4010 TTTTCTCTCAACCTCTTCCACCATATCAGTCTTAAGAAGATATCTTTCTCAGGTTTCAGAC 3951

QY 180 TAATATAACAGAGTATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCATCTAGATG 239
DB 3950 CAATTAGCGGAGTCGACACCGCTGTCATATAAAGCTTCAAGGGTGCCATCTTTGATA 3891

QY 240 CCAGATTGAAAAGTGTATGTAGGCGAACTCAACTAAGCGTAAATAATCTCTCAACTA 299
DB 3890 CTTTCTTGATAACTATTATTATGAAAACTCCGCTTAAAGGTAAACATTCGTCCTCCATTTT 3831

QY 300 CCTTAGTAATAATACATAGCTCCAAATCGTATCTCTAGTATATGAATCACTTCTCTCA 359
DB 3830 AGTGAGAAATCCAGGACACATGCCCGTAGCATATCTTCAAGTATTTGGTTCACTCTCTCA 3771

QY 360 AATTGACCATCGTCTGAGGATGGAATGAGACCGGTGCCACCGATT 406
DB 3770 GTTTGCTCACTGCTTTGAGGATGATAGTTCGAACCTGTGGAGCAGTTT 3724
```

RESULT 9

```
US-10-437-963-78436/c
; Sequence 78436, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 78436
; LENGTH: 2373
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2373)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_78243C.1
US-10-437-963-78436

Query Match          9.0%; Score 129.2; DB 19; Length 2373;
Best Local Similarity 60.1%; Pred. No. 7e-09;
Matches 215; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 38 CTTGATATCTATGATTTTCAGATTTTCATAGAAGACTTCTATCTATCAGAGACCGCTCAG 97
DB 2166 CTCCTAATTCCTGCTTTTGACATCGGCGTAGCTCTTCTGCTACTTTTGAGCTGCAAGCAT 2107

QY 98 AGGATCCCAAAATTAGTCTAAAATTATCTTCAGTCTCGGAAACCAACTCAGGACCCCAAAAC 157
DB 2106 GTGTTGTGATTTCTGCTATCCGTTCTGTGGTCTGTGTGACCAATCGGCTCTCTAAAT 2047

QY 158 CCGTCGCTCACCAACTCAGTCTTAATATAACAGAGTATGACACTTATGACCATATAGAGC 217
DB 2046 AGCTCTCTCCCTATGGAATCCCAACACATATAGGATCGACACTTTCTTCCATAAAGTGC 1987

QY 218 CTCGTAGGTGCCATCTAGATGCCAGATTGGAACCTGTTATTGTAGGCGAACTCAACTAA 277
```

Db 1986 CTCGTGTTGCCACTTGAATGCTAGCATGGAACACTATTGTTGAAGTAAATTTCTACAAG 1927
Qy 278 CGGTAAATAATCTCTCACTACCTTAGTAAATACATAGCTCCAAATCGTATCCTC 337
Db 1926 CGGCAATATGATCTCTCCCAACTGCGCTTTCCATGAAGGATACATGATCTCAACATATCCTC 1867
Qy 338 TAGTATATGAATCACTCTTCAAAATGACCATCGGTCTGAGGATGGAATGACAGCCGG 395
Db 1866 TAAAGTTTGAATAGTCGCTCAGATTGACCATGACGCTGAGGGTGAAAAGCCGTACTG 1809

RESULT 10

US-10-437-963-78434/c
; Sequence 78434, Application US/10437963
; Publication No. US20040123343A1

GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 78434
; LENGTH: 4641
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_78241C.1

US-10-437-963-78434

Query Match 9.0%; Score 129.2; DB 19; Length 4641;
Best Local Similarity 60.1%; Pred. No. 8.2e-09;
Matches 215; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy 38 CTTGATATCTATGATTTTTCAGATTGCAATAGACTTCTATCTATCATCAGAGACGCTGCAG 97
Db 4032 CTCATAATCTCGTCTTGACATCGGCGTAGCTCTTCGCTACTCTTCGAGCTGCAAGCAT 3973
Qy 98 AGGATCCCAATATTAGTCTAAATTTATCTTCAGTCTCGGAAACCAACTCAGGACCCAAAC 157
Db 3972 GTGTTGTCGATTTCTACTATCCGTTCTGTGTCGTTGTACCCCAATCGGGTCTCTAAAT 3913
Qy 158 CCGTCGCTCACCCTCACTCACTATATAACAGAGATGACACTTATGACCATATAGAGC 217
Db 3912 AGTTCTCTCCGCTATGGAATCCCAACACAAATGGGATCGACACTTTCTTCATAAGTGC 3853
Qy 218 CTCGTAGTCCCTCTAGATGCGGAGTTGGAATCTGTTATTGTAGGCGCACTCAACTAA 277
Db 3852 CTCATATGTCGCTTTGAATGCTAGTGAAGAACTATTGTTGAAGCAATTTCTACAG 3793
Qy 278 CGGTAAATAATCTCTCACTACCTTAGTAAATACATAGCTCCAAATCGTATCCTC 337
Db 3792 CGGCAATATGATCTCTCCCAACTGCGCTTCCATGAAGGATACATGATCTCAACATATCCTC 3733
Qy 338 TAGTATATGAATCACTTTCTCAAAATGACCATCGGTCTGAGGATGGAATGACAGCCGG 395
Db 3732 TAAAGTTTGAATAGTCCGCTCAGATTGACCATGACGCTGAGGGTGAAAAGCCGTACTG 3675

RESULT 11

US-10-425-115-177419/c
; Sequence 177419, Application US/10425115
; Publication No. US2004021472A1

GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 177419
; LENGTH: 4884
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_31023C.1

Query Match 9.0%; Score 128.8; DB 20; Length 4884;
Best Local Similarity 61.2%; Pred. No. 9.5e-09;
Matches 208; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 177419
; LENGTH: 4884
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_93395C.1

US-10-425-115-177419

Query Match 9.0%; Score 128.8; DB 20; Length 1155;
Best Local Similarity 61.2%; Pred. No. 6.7e-09;
Matches 208; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

Qy 58 GATTTCATAGACTTCTATCTATCAGAAGAGCGCTCAGAGGATCCCAAAATAGTCTAA 117
Db 697 GGTCTGCATAACTCTTTTCCGAGTCTGAGCTTCTTCATATATGATATATTTCTGAA 638
Qy 118 AATATCTTTCAGTCTCGGAAACCAACTCAGGACCCAAACCCGCTCAGCCCAACTCAG 177
Db 637 CTTTCTTCGACCTCTTTTACCATAATCAGGCTTAAGAGTACTTCTCCAGGTTTCCAG 578
Qy 178 TCTAATATAACAGAGTATGACATTATGACCATATAGAGCTCTGTAAGTGCCATCTAGA 237
Db 577 ACCAATTTAGCGGAGTTTCGACACCGTTGCCATACAAAGCTTCAAAAGGTGCCATCTTGA 518
Qy 238 TGCAGATTGGAACCTGTTATTGTAGGCGCACTCAACTAAGCGTAAATAAATCTCTCAAC 297
Db 517 TACTTTCTTGATTAACCTGTTATTGTATGAAACTCCGCTAAGGGCAACATTCATCCCAT 458
Qy 298 TACTTATGATTAATAATCACATAGCTCCAAATCGTATCTCTAGTATATGATATCACTTCT 357
Db 457 TCTGTGAGAGTCCGAAACACATGCCCGCAGCATATCTTCAGTATTTGATTGACCTCT 398
Qy 358 CAAATTCACCATCGGTCTGAGGATGGAATGACAGCCGGTG 397
Db 397 CAGTCTGCCACTGTTTGGAGGATGATAAGCCGAACCTGTG 358

RESULT 12

US-10-425-115-109131/c
; Sequence 109131, Application US/10425115
; Publication No. US2004021472A1

GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 109131
; LENGTH: 4884
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_31023C.1

US-10-425-115-109131

Query Match 9.0%; Score 128.8; DB 20; Length 4884;
Best Local Similarity 61.2%; Pred. No. 9.5e-09;
Matches 208; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

Qy	58	GATTTCATAAGACTTCTTATCTATCAGAAGACGGCTCAGAGGATCCCAATTAGTCTAA	117
Db	4426	GGTCGCATAACTCTTTTGTGCGAGTCGAGCTTCTTCATATTATGAATAATTTATGAA	4367
Qy	118	AATTATCTTCAGTCTCGGAAACCAACTCAGAGCCCAAAACCGTCGCTCACCNACTCAG	177
Db	4366	CCTTTTCTTCGACCTCTTCCACCATATCAGGCTTAAAGAATGATCTTTTCTCCAGGTTTCAG	4307
Qy	178	TCTAATATAACAGAGTAGACACTTATGACCATATAGAGCCTCGTAAAGTGCCACTTAGA	237
Db	4306	ACCAATTTAGCGGAGTTCGACACCGTCGCCCATATAAAGCTTCAAAAGGTGCCACTTTGA	4247
Qy	238	TGCAGATTGGAACCTGTTATTGTAGGGAACTCACTAACGGTAAAAAATCCTCTCAAC	297
Db	4246	TACTTTCTTGATAACTGTATTGTATGAAAACCTCCGCTAGGGGCAAACTTCATCCCAAT	4187
Qy	298	TACCTTAGTATAAATCACATAGCTCCAAATCGTATCCTCTAGTATATGAATCACCTTCT	357
Db	4186	TCTGTGAGAGTTCAGAACACATGCCCGCAGCATATCTCGAGTATCTGATTTACCCCTCT	4127
Qy	358	CAAAATGACCATCCGCTGAGGATGGAATGCAGACCGGTG	397
Db	4126	CAGTCTGTCCACTTGTGTTTGAGGATGATAGGCCGAACCTGTG	4087

RESULT 13

```

US-10-425-115-177411
* Sequence 177411, Application US/10425115
* Publication No. US20040214272A1
* GENERAL INFORMATION:
* APPLICANT: La Rosa, Thomas J.
* APPLICANT: Kovalic, David K.
* APPLICANT: Zhou, Yihua
* APPLICANT: Cao, Yongwei
* TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
* TITLE OF INVENTION: Plants
* FILE REFERENCE: 38-21(53222)B
* CURRENT APPLICATION NUMBER: US/10/425,115
* CURRENT FILING DATE: 2003-04-28
* NUMBER OF SEQ ID NOS: 369326
* SEQ ID NO 177411
* LENGTH: 5433
* TYPE: DNA
* ORGANISM: Zea mays
* FEATURE:
* NAME/KEY: unsure
* LOCATION: (1)..(5433)
* OTHER INFORMATION: unsure at all n locations
* FEATURE:
* OTHER INFORMATION: Clone ID: MFT4577_93388C.1
US-10-425-115-177411

```

```

QY 300 CCTTAGTAAATAACATAGCTCCAAATCGTATCTCTAGTATATATGAATCACCTTCTCA 359
Db 701 AGTGAGAAATCAGAGACATGCGCGTAGCATATCTTCAAGTATTTGGTTACCCCTCTCA 760
QY 360 AATTGACCATCGGTCGTGAGGATGGAATGCAGACCCGGTG 397
Db 761 GTCTGCCCACTGGTTTGAGGATGATAGCCGAACTGTG 798

RESULT 14
US-10-298-122-2/c
; Sequence 2, Application US/10298122
; Publication No. US20030221214A1
; GENERAL INFORMATION:
; APPLICANT: Gmitter, Frederick G
; APPLICANT: Deng, Zhanao
; APPLICANT: Zhang, Hongbin
; TITLE OF INVENTION: CITRUS TRISTEZA VIRUS AND METHODS OF USE
; FILE REFERENCE: 5853-220
; CURRENT APPLICATION NUMBER: US/10/298,122
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 58965
; TYPE: DNA
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Poncirus trifoliata and Citrus grandis hybrid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2145)..(2320)
; OTHER INFORMATION: n denotes unsequenced nucleotides
US-10-298-122-2

```

RESULT 15

US-10-437-963-78438/c
; Sequence 78438, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 78438
 ; LENGTH: 4641
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_78245C.1
 US-10-437-963-78438

Query Match 8.8%; Score 126; DB 19; Length 4641;
 Best Local Similarity 59.5%; Pred. No. 2.3e-08;
 Matches 213; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy	38	CTTCATATCTATGATTTTCAGATTTCATAGACTTCTATCTATCAGAGACGCTGCAG	97
Db	4161	CTCTAATTCTCGTCTTCTGACATCGGCGTAGCTCTTCTGCTACTTTGAGCTGCAAGCAT	4102
Qy	98	AGGATCCCAAAATTAGTCTAAAAATTATCTTCAGTCTCGGAAACCAACTCAGGACCCAAAAC	157
Db	4101	GTGTTGTCGTATTTCTGCTATCCGTTCTGTGGTCTGTGTACCCACCGGGTCTTAAAT	4042
Qy	158	CCGTCGCTCACCCAACTCAGTCTTAATATAACAGAGATGACACTTTATGACCATATAGAC	217
Db	4041	AGCTCTCTCCCTATGGAATCCCAACACAAATGGGGATCGACACTTTCTTCCATAAAGTGC	3982
Qy	218	CTCGTAAGTGCCATCTAGATGCCAGATTGGAACTGTTATTTAGGGGAACTCAACTAA	277
Db	3981	CTCATATGGTGCCATTTGAATGTCTAGCATGGAACTATTTGTGAAGAAAATTTCTCAAG	3922
Qy	278	CGGTAAAAAATCCCTCTCAACTACCTTAGTATAAATCACATAGCTCCAAATCGTATCCTC	337
Db	3921	CGGCAATATGATCCCTCCCACTGCCCTTTCCATGAAAGGATACATGATCTCAACATATCCTC	3862
Qy	338	TAGTATATGAATCACCTTTCTCAAAATGACCATCGGTCTGAGGATGGAATGCAGACCGG	395
Db	3861	TAAAGTTTGAATAGTCCGCTCAGATTGACCATCAGTCTGAGGGTGAAGGCGGTACTG	3804

Search completed: August 13, 2005, 22:17:26
 Job time : 8245 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 12, 2005, 18:33:08 ; Search time 1232 Seconds

(without alignments)
6885.543 Million cell updates/sec

Title: US-10-089-543-2

Perfect score: 1433

Sequence: 1 actataggcagcgggtgc.....tgagagaaatccttcacatc 1433

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

1: geneseqn1980s:.*
2: geneseqn1980s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1433	100.0	1433	6	ABK15699 Cotton fi
2	120	8.4	2000	8	ADA71742 Rice gene
3	113.8	7.9	31491	4	ADL10203 Pepper Bs
4	113.8	7.9	31491	6	AAF63301 Pepper Bs
5	113.8	7.9	31491	6	AAK98863 Nucleic a
6	107.4	7.5	8056	8	ABZ10246 Haematopo
7	105.8	7.4	8056	8	ABZ10100 Haematopo
8	105	7.3	8222	8	ACF62816 Colon can
9	105	7.3	11222	10	ADB54318 Pretreat
10	105	7.3	9590	3	AAF22281 BAC conta
11	104.8	7.3	158001	12	ADL17884 Human pho
12	104.4	7.3	4985	6	ABQ75107 Anopheles
13	104.4	7.3	4985	10	ACF79720 Mosquito
14	103.8	7.2	163319	3	AAF22306 Arabidops
15	103.4	7.2	8222	8	ACF62794 Colon can
16	103.4	7.2	11222	10	ADB54190 Pretreat
17	102.4	7.1	2000	12	ADJ41131 Plant cdn
18	102.2	7.1	109973	3	AAF22298 BAC conta
19	101.6	7.1	8056	8	ABZ10246 Haematopo
20	101	7.0	110000	3	AAF22305_08 Continuation (9 of

C	21	100	7.0	8056	8	ABZ10100 Haematopo
C	22	99.2	6.9	79122	3	AAF22294 BAC conta
C	23	99	6.9	101786	3	AAF22293 BAC conta
C	24	97.6	6.8	86584	3	AAF22292 BAC conta
C	25	97.4	6.8	110000	3	AAF22303_2 Continuation (3 of
C	26	96.2	6.7	15548	6	ABL34155 Human imm
C	27	96.2	6.7	92584	3	AAF22288 BAC conta
C	28	96	6.7	18154	6	ABL32254 Human imm
C	29	95.4	6.7	6109	6	ABL32326 Human imm
C	30	95.4	6.7	6109	6	AA61077 Human gen
C	31	94.6	6.6	6963	6	ABL32979 Human imm
C	32	94.4	6.6	419	8	ABX46069 Bovine ES
C	33	94.4	6.6	5286	13	ADS89278 Ades9278 Oligonuc
C	34	94.4	6.6	5286	13	ADS89552 Oligonuc
C	35	93.6	6.5	9642	6	ABL32357 Human imm
C	36	91.8	6.4	6216	6	ABK39932 Human che
C	37	91.8	6.4	6216	6	ABL70139 Chemical
C	38	91.8	6.4	7167	6	ABL32400 Human imm
C	39	91.2	6.4	6048	6	ABQ67002 Human ang
C	40	91.2	6.4	12237	6	ABL34358 Human imm
C	41	90.8	6.3	7814	4	AA646530 Tumour su
C	42	90.6	6.3	6419	6	ABL32267 Human imm
C	43	90.2	6.3	778	6	ABQ15588 Oligonuc
C	44	90.2	6.3	778	6	ABQ15589 Oligonuc
C	45	90.2	6.3	6175	6	ABL33307 Human imm

ALIGNMENTS

RESULT 1

ABK15699

ID ABK15699 standard; DNA; 1433 BP.

XX

AC ABK15699;

XX 21-MAY-2002 (first entry)

XX Cotton fibre-specific beta tubulin, CFTUB2, promoter fragment.

XX Cotton; ds; CFTUB2; fibre-specific beta tubulin; transgenic. promoter;
KW plant; anthocyanin gene; silk protein gene; cotton fibre strength;
KW polyhydroxybutyrate.

XX Gossypium hirsutum.

XX Key Location/Qualifiers
promoter 449..1433
/*tag= a

FT FT /note= "Cotton fibre-specific promoter. This sequence is
FT specifically claimed in claim 3"

XX WO200210377-A1.

XX 07-FEB-2002.

XX 01-AUG-2000; 2000WO-SG000111.

XX 01-AUG-2000; 2000WO-SG000111.

XX (MOLE-) INST MOLECULAR AGROBIOLOGY.

XX Cai L, Li X, Cheng N, Liu J;

PI WPI; 2002-217120/27.

XX New fiber-specific beta-tubulin promoter from cotton for controlling gene
PT expression in cotton fibers and creating transgenic plants, in particular
PT cotton plants, having altered fiber characteristics.

XX Claim 2; Fig 2; 30pp; English.

XX The invention relates to a promoter that is cotton fibre-specific,

CC comprising the promoter of the cotton beta-tubulin gene CFTUB2. The
 CC promoter is fibre-specific in cotton and controls specific gene
 CC expression at the transcriptional level in cotton fibres and is useful
 CC for improving cotton fibres to create new cotton varieties with high
 CC fibre quality and yield by gene manipulation. The promoter is useful for
 CC creating transgenic plants, in particular cotton having altered fibre
 CC characteristics, and permits selective expression of a transgene in the
 CC cotton fibre, permitting greater latitude in the types of transgenes
 CC employed. Examples of expression of desirable genes in cotton fibres, but
 CC not in other parts of the cotton plants, include anthocyanin genes for
 CC coloured cotton, silk protein genes from silk worm or spiders for
 CC increased strength of cotton fibre, and biosynthesis of
 CC polyhydroxybutyrate in cotton fibre for improved thermal properties and
 CC insulating characteristics. The promoter can improve cotton fibres to
 CC create new cotton varieties with higher fibre quality and yield. The
 CC present sequence is the CFTUB2 promoter fragment
 XX
 SQ Sequence 1433 BP; 466 A; 285 C; 180 G; 502 T; 0 U; 0 Other;

Query Match 100.0%; Score 1433; DB 6; Length 1433;
 Best Local Similarity 100.0%; Pred. No. 8.1e-189;
 Matches 1433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTATAGGGCAGCGGTGTCGACGCCCGGGCTGGTCTTGATATCATGATTTTCAGAT 60
 DB 1 ACTATAGGGCAGCGGTGTCGACGCCCGGGCTGGTCTTGATATCATGATTTTCAGAT 60
 QY 61 TTGCATAAGACTTCTATCTATCATGAGAGACGCCCTGCGAGAGGATCCCAATTAGTCTAAAT 120
 DB 61 TTGCATAAGACTTCTATCTATCATGAGAGACGCCCTGCGAGAGGATCCCAATTAGTCTAAAT 120
 QY 121 TATCTTCAGTCTCGGAAACCAACTCAGGACCCCAAAACCGGTGCTCACCACCACTCAGTCT 180
 DB 121 TATCTTCAGTCTCGGAAACCAACTCAGGACCCCAAAACCGGTGCTCACCACCACTCAGTCT 180
 QY 181 AATATAACAGAGATGACACTTATGACCATATAGAGCCTCGTAGGTGCCATCTAGATGC 240
 DB 181 AATATAACAGAGATGACACTTATGACCATATAGAGCCTCGTAGGTGCCATCTAGATGC 240
 QY 241 CAGATTGGAAACGTATTATGTCGCGAATCACTAAGCGGTAAAGAAATCCTCTCAACTAC 300
 DB 241 CAGATTGGAAACGTATTATGTCGCGAATCACTAAGCGGTAAAGAAATCCTCTCAACTAC 300
 QY 301 CTTAGTAAATAAATCACATAGCTCCTATCTCTAGTATATGAATCACCCTCTCTCAA 360
 DB 301 CTTAGTAAATAAATCACATAGCTCCTATCTCTAGTATATGAATCACCCTCTCTCAA 360
 QY 361 ATTGACCATCGGTCTGAGGATGGAATGACAGCCGGTGCCACCGATTCTAATGGTACCT 420
 DB 361 ATTGACCATCGGTCTGAGGATGGAATGACAGCCGGTGCCACCGATTCTAATGGTACCT 420
 QY 421 ATAAAAAATTTATTTTAAAAAATGATGTGACAGGTGGTGGAGAGAGGTCTAC 480
 DB 421 ATAAAAAATTTATTTTAAAAAATTTGATGTGACAGGTGGTGGAGAGAGGTCTAC 480
 QY 481 CGATTGTCAGGTGCGCAACCAATTTTATTTTACCTCCTCGCTAGATTTCGTAATACTAT 540
 DB 481 CGATTGTCAGGTGCGCAACCAATTTTATTTTACCTCCTCGCTAGATTTCGTAATACTAT 540
 QY 541 TGCATTATCTCAATTCATTATTTAATAATTTTATATTTTATATTTTATATTTTATATTTT 600
 DB 541 TGCATTATCTCAATTCATTATTTAATAATTTTATATTTTATATTTTATATTTTATATTTT 600
 QY 601 ATACTTTACTTTTTTAAAAAGAAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 660
 DB 601 ATACTTTACTTTTTTAAAAAGAAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
 QY 661 TCTAATACTTTTACTTTTTTAAAAAGAAATTTCAATTTGCGTTTCTTTTAAATTTAGTTT 720
 DB 661 TCTAATACTTTTACTTTTTTAAAAAGAAATTTCAATTTGCGTTTCTTTTAAATTTAGTTT 720
 QY 721 TAAATTCATATAATAAATAATTCATGATCGGATAGTGTGGTCAAAAGTCAAGTCAAC 780

DB 721 TAATTCATATAATAAATAATTCATGATCGGATAGTGTGGTCAAAAGTCAAGTCAAC 780
 QY 781 ATGAATTTTGTGGAGAAAAATAAAAAATTAACACATTTTTCGATTAAATTTTATATATA 840
 DB 781 ATGAATTTTGTGGAGAAAAATAAAAAATTAACACATTTTTCGATTAAATTTTATATATA 840
 QY 841 TATAATAATATAAACACATTTTATTTAATGTGTCAATAATTTTAAATTTTAAATTTT 900
 DB 841 TATAATAATATAAACACATTTTATTTAATGTGTCAATAATTTTAAATTTTAAATTTT 900
 QY 901 CAGCACAACAATTACACTCTCATTAATTTTAAATTTTACCATTAATTTTAAATTTT 960
 DB 901 CAGCACAACAATTACACTCTCATTAATTTTAAATTTTACCATTAATTTTAAATTTT 960
 QY 961 AGGACAATTTATTTTAAATTTTCAACCTCCATTAATGATATTTTATTTTGTTCGATA 1020
 DB 961 AGGACAATTTATTTTAAATTTTCAACCTCCATTAATGATATTTTATTTTGTTCGATA 1020
 QY 1021 CTTCTTATTTCACTCCTAAACATTAATCAATTAACCCCAATTTTGAATTTTCTTA 1080
 DB 1021 CTTCTTATTTCACTCCTAAACATTAATCAATTAACCCCAATTTTGAATTTTCTTA 1080
 QY 1081 ACTTATTTCAATTTGGCTCTGGTCCATCTGGAAGGCCACCGTCCAGGCTGTCCAAC 1140
 DB 1081 ACTTATTTCAATTTGGCTCTGGTCCATCTGGAAGGCCACCGTCCAGGCTGTCCAAC 1140
 QY 1141 CACACTTTGCCAGCTCATCAATTTCCAGTAACTACATTTGTTACAGTTTACTAAGCAATCCC 1200
 DB 1141 CACACTTTGCCAGCTCATCAATTTCCAGTAACTACATTTGTTACAGTTTACTAAGCAATCCC 1200
 QY 1201 AATTTTCAAAAAATTTCAATTTTCCAGGAAAAACGAAACGTCGTTACTAACCGACCTAAAC 1260
 DB 1201 AATTTTCAAAAAATTTCAATTTTCCAGGAAAAACGAAACGTCGTTACTAACCGACCTAAAC 1260
 QY 1261 CAGCTCAACCTGCGTCAATTTAAACGGAATCTTTTAACTCCTCTATATAACCCAAACCA 1320
 DB 1261 CAGCTCAACCTGCGTCAATTTAAACGGAATCTTTTAACTCCTCTATATAACCCAAACCA 1320
 QY 1321 CTTCTCATCACTTTTCCCATAAAAAGAAATTTCCGGAATTTCTTATTTCTTTTATATTTT 1380
 DB 1321 CTTCTCATCACTTTTCCCATAAAAAGAAATTTCCGGAATTTCTTATTTCTTTTATATTTT 1380
 QY 1381 CCTCTCAATTTCCCGTCACTTTCCGAGAAAAATGAGAAAAATCCTTTCATC 1433
 DB 1381 CCTCTCAATTTCCCGTCACTTTCCGAGAAAAATGAGAAAAATCCTTTCATC 1433

RESULT 2

ADA711742
 ID ADA71742 standard; DNA; 2000 BP.

XX ADA71742;

XX AC

XX 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 5067.

XX KW Plant; bacterial infection; fungal infection; viral infection; rice;

XX KW gene; ds.

XX OS Oryza sativa.

XX XX WO200300898-A1.

XX PN 03-JAN-2003.

XX PD 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

```

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX Claim 27; SEQ ID NO 5067; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC the expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
XX Sequence 2000 BP; 569 A; 433 C; 409 G; 589 T; 0 U; 0 Other;
SQ
Query Match 8.4%; Score 120; DB 8; Length 2000;
Best Local Similarity 60.4%; Pred. No. 4.9e-08;
Matches 198; Conservative 0; Mismatches 130; Indels 0; Gaps 0;
QY 63 GCATAAGACTTCTATCTATCAGAGACGCGCTGAGAGGATCCCAAATTAGTCTAAATTA 122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
421 GCGTAATCTCTGACGGTTCTGTGTGTCTTCAACGTTCTCGTATCAGCTTAACCTGT 480
-QY 123 TCCTCAGTCTCGGAACCACTCAGGACCCAAACCGTCGCTCACCACCACTCAGTCTAA 182
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
481 TCCTCTGCTGACTTGTAGTATGTGAGGACCGAATATCAGCGCTTCGCCCACTTCATCCAG 540
QY 183 TATAACAGAGTATGACACTTATGACATATAGAGCCTCGTAAGGTGCGCATCTAGATGCCA 242
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
541 CATNAGGGAGTGGCGACTTCTTCCGACATCTGCTTCGCGGACATCTGATGCTG 600
QY 243 GATTGGAACCTGTTATTGTAGGCGAATCACTAACCGTAAATAATCCTCTCAACTACCT 302
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
601 GCCTGATAGCTGTTGTTGTATGAGAACTCAGCATACGCGCAACCAACGATCCCAAGTCCCT 660
QY 303 TAGTAATAATACATAGTCCAAATCGTATCCTCTAGTATATGAAATCACTTCTCAAT 362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
661 TCGAAATCCAGGCGACACGACGATGATGTTCTTAGGATTTGGTTTACCCCTTCTGTC 720
QY 363 TGACCATCGGTCTGAGGATGGAATGCCAG 390
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
721 TGACCATCGGTCTGCGGATGTTAGGCCG 748
RESULT 3
AAD10203
ID AAD10203 standard; DNA; 31491 BP.
XX
XX AAD10203;
XX
XX 24-SEP-2001 (first entry)
XX
XX Pepper Bs2 gene.
XX
XX Bs2 gene; pepper; resistance gene; plant pathogen; virulence gene;
KW bacterial spot disease; Xanthomonas campestris pv. vesicatoria; AvrBs2;
KW hypersensitive response; transgenic plant; tomato; tobacco; rice; corn;
KW wheat; ds.
XX
XX Capsicum annum.
XX
XX Key Location/Qualifiers
FH 1.502
FT promoter
FT /*tag= a

```

```

FT exon 503..554
FT /*tag= b
FT /number= 1
FT /note= "This region contains a portion of 5' untranslated
FT region (5' UTR)"
FT intron 555..1439
FT /*tag= c
FT /number= 1
FT /note= "The 5' untranslated region (5' UTR) continues in
FT this region"
FT exon 1440..4162
FT /*tag= d
FT /number= 2
FT /note= "The region 1440-1479 contains 5' UTR which is
FT followed by the coding region"
FT CDS 1480..31219
FT /*tag= e
FT /product= "Bs2 protein"
FT intron 4163..31184
FT /*tag= f
FT exon 31185..31216
FT /*tag= g
FT /number= 3
FT 3'UTR 31220..31491
FT /*tag= h
XX US6262343-B1.
XX 17-JUL-2001.
XX 23-JUL-1999; 99US-00360186.
XX 23-JUL-1998; 98US-0093957P.
XX (REGC ) UNIV CALIFORNIA.
XX Skaskawicz BJ, Dahlbeck D, Tai TH;
XX WPI; 2001-450496/48.
XX P-PSDB; AAE05409.
XX Nucleic acid molecules encoding Bs2 protein, useful for producing
XX transgenic plants having resistance to the plant pathogen Xanthomonas
XX campestris.
XX Claim 7; Col 21-50; 37pp; English.
XX
XX The present sequence is Bs2 gene from pepper. The Bs2 gene is shown to
XX confer resistance to plant pathogen Xanthomonas campestris pv.
XX vesicatoria which causes bacterial spot disease. The Bs2 protein has Bs2
XX biological activity, i.e., when co-expressed in a plant with a X.
XX campestris AvrBs2 gene product, it produces a localised hypersensitive
XX response. The protein includes a nucleotide binding motif and leucine
XX rich repeats of the type found in other plant resistance genes. The Bs2
XX nucleic acid molecule is useful for producing transgenic plants such as
XX pepper, tomato, tobacco, broccoli, cauliflower, cabbage, cowpea, grape,
XX canola, bean, soybean, rice, corn, wheat, barley, citrus, cotton, cassava
XX and walnut, that are resistant to plant pathogen Xanthomonas campestris.
XX The transgenic plants produced using Bs2 molecule develop a
XX hypersensitive response to the pathogen at the site of inoculation and
XX show an enhanced resistance to systemic infection. The Bs2 nucleic acid
XX molecule is also useful as polymerase chain reaction (PCR) primers for
XX amplifying portions of Bs2 nucleic acid molecule, as sequencing primers
XX to verify the authenticity of an amplified molecule, and as hybridisation
XX probes
XX
XX Sequence 31491 BP; 9652 A; 6344 C; 5992 G; 9503 T; 0 U; 0 Other;
SQ
Query Match 7.9%; Score 113.8; DB 4; Length 31491;
Best Local Similarity 59.5%; Pred. No. 2.7e-07;
Matches 210; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

```

QY 38 CTTGATATCTATGATTTTCAGATTTTGCATAAGACTTCTATCTATCATCAGAGAGCGCCTGCGAG 97
DB 22086 CTTCTAGTCTCTTCTCCGAAACATCGGCATAGGACATCTAAGGACTTTTGAGCAGTCTTGAC 22145
QY 98 AGGATCCCAAAATAGTCTAAAATATCTTCAGTCTCGGAAACCAACTCAGGACCCCAAAAC 157
DB 22146 TCG-TCTCTAATGATTTTCACTTTCTCCATCGCCTTATGAACAAGATTAGGCCCATACAA 22204
QY 158 CGTCTGCTCACCACCACTCAGTCTAATATAACAGAGTATGACACTTATGACCATATATAGAGC 217
DB 22205 CTGAGTCTCACCACCTTCATACATCTATCGGAGACCTTATATCTCTCTCAATATCAAAGC 22264
QY 218 CTGCTAAGGTGCCATCTAGATGCCAGATTGGAACTGTGTTATTGTAGCGCAACTCAACTAA 277
DB 22265 CTCAAAAGGAGCCATCTTGATGCTGGCATGGTAGTTATTATTGTAAGCGCAATTCACACCG 22324
QY 278 CGGTAAAAAATCTCTCAACTACTCTTAGTAATAAATCAACATAGCTCAAAATCGTATCTTC 337
DB 22325 TGGCAGGTGATCTACCCAACTACCTTTTGAATCAATTAACGATGCGCCTAAACATATCTTC 22384
QY 338 TAGTATATGAATCACCTTCTCAAAATTGACCATCGGTCTGAGGATGGAATGCAG 390
DB 22385 GAGGGTCTGAATGGTAGCTCAGCTTGTCATCCGCTGAGGGTGAACCTG 22437

RESULT 4
AAAF63301
ID AAF63301 standard; DNA; 31491 BP.
XX
AC AAF63301;
XX
DT 02-MAY-2001 (first entry)
XX
DE Pepper Bs2 resistance gene.
XX
KW Bs2; pepper; pathogen resistant; Xanthomonas campestris pv vesicatoria;
KW Xcv; bacterial spot disease; transgenic plant; crop; fruit; flower; ds.
XX
OS Capsicum annum.
XX
FN W0200107635-A1.
XX
PD 01-FEB-2001.
XX
PF 23-DEC-1999; 99WO-US030891.
XX
PR 23-JUL-1999; 99US-00360186.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Staskawicz BJ, Dahlbeck D, Tai TH;
XX
DR WPI; 2001-168560/17.
DR P-PSDB; AAB72198.
XX
PT Novel Bs2 polypeptide from Capsicum annum for producing transgenic
PT plants having resistance to bacterial spot disease caused by Xanthomonas
PT campestris pv vesicatoria (Xcv).
XX
PS Claim 3; Page 40-57; 72pp; English.

XX This invention relates to the amino acid sequence of the Bs2 protein
XX isolated from Capsicum annum (pepper). The protein and DNA sequences of
XX Bs2 can be used to confer resistance to the plant pathogen Xanthomonas
XX campestris pv vesicatoria (Xcv) which causes bacterial spot disease. Bs2
XX DNA and protein sequences are useful for producing transgenic plants such
XX as pepper, tomato, tobacco, broccoli, cauliflower, cabbage, canola,
XX cowpea, grape, bean soybean, rice, corn, wheat, barley, citrus, cotton,
XX cassava and walnut, having resistance to X. campestris. The protein and
XX DNA molecule are also useful for producing transgenic alfalfa, flax,
XX sunflower, safflower, brassica, peanut, clover, lettuce, cucurbits,
XX potato, carrot, radish, pea, lentils, apples, pears, peaches, apricots,
XX carnations and roses having resistance to X. campestris. The present

CC sequence represents the Bs2 gene of the invention
XX
SQ Sequence 31491 BP; 9652 A; 6345 C; 5992 G; 9502 T; 0 U; 0 Other;
Query Match 7.9%; Score 113.8; DB 4; Length 31491;
Best Local Similarity 59.5%; Pred. No. 2.7e-07;
Matches 210; Conservative 0; Mismatches 142; Indels 1; Gaps 1;
QY 38 CTTGATATCTATGATTTTTCAGATTTTGCATAAGACTTCTATCTATCATCAGAGAGCGCCTGCGAG 97
DB 22086 CTTCTAGTCTCTTCTCCGAAACATCGGCATAGGACATCTAAGGACTTTTGAGCAGTCTTGAC 22145
QY 98 AGGATCCCAAAATAGTCTAAAATATCTTCAGTCTCGGAAACCAACTCAGGACCCCAAAAC 157
DB 22146 TCG-TCTCTAATGATTTTCACTTTCTCCATCGCCTTATGAACAAGATTAGGCCCATACAA 22204
QY 158 CGTCTGCTCACCACCACTCAGTCTAATATAACAGAGTATGACACTTATGACCATATATAGAGC 217
DB 22205 CTGAGTCTCACCACCTTCATACATCTATCGGAGACCTTATATCTCTCTCAATATCAAAGC 22264
QY 218 CTGCTAAGGTGCCATCTAGATGCCAGATTGGAACTGTGTTATTGTAGCGCAACTCAACTAA 277
DB 22265 CTCAAAAGGAGCCATCTTGATGCTGGCATGGTAGTTATTATTGTAAGCGCAATTCACACCG 22324
QY 278 CGGTAAAAAATCTCTCAACTACTCTTAGTAATAAATCAACATAGCTCAAAATCGTATCTTC 337
DB 22325 TGGCAGGTGATCTACCCAACTACCTTTTGAATCAATTAACGATGCGCCTAAACATATCTTC 22384
QY 338 TAGTATATGAATCACCTTCTCAAAATTGACCATCGGTCTGAGGATGGAATGCAG 390
DB 22385 GAGGGTCTGAATGGTAGCTCAGCTTGTCATCCGCTGAGGGTGAACCTG 22437

RESULT 5
AAK98863
ID AAK98863 standard; DNA; 31491 BP.
XX
AC AAK98863;
XX
DT 17-MAY-2002 (first entry)
XX
DE Nucleic acid of the pepper Bs2 gene.
XX
KW Resistance; plant; pathogen; Xanthomonas campestris; Capsicum annum;
KW pepper; Xcv chromosome; Bs2; gene; ds; transgenic plant.
XX
OS Capsicum annum.
XX
FH Key Location/Qualifiers
FT promoter 1..502
FT /*tag= b
FT /note= "Promoter region"
FT exon 503..554
FT /*tag= c
FT /number= 1
FT /note= "Exon 1 (5' UTR)"
FT intron 555..1439
FT /*tag= d
FT /number= 1
FT exon 1440..1479
FT /*tag= e
FT /number= 2
FT /note= "Exon 2 (5' UTR (continued))"
FT CDS 1480..31219
FT /*tag= a
FT /product= "Protein of pepper Bs2"
FT /note= "This CDS contains 2 introns"
FT exon 1480..4162
FT /*tag= f
FT /number= 2
FT /note= "Exon 2 continued"
FT intron 4163..31184
FT /*tag= g

[illegible]

XX The invention relates to a novel method for detecting and differentiating
CC between colon cell proliferative disorders associated with at least one
CC gene or its regulatory regions. The method comprises contacting a target
CC nucleic acid in a biological sample obtained from the subject with at
CC least one reagent or a series of reagents, where the reagent or series of
CC reagents, distinguishes between methylated and non methylated CpG
CC dinucleotides within the target nucleic acid. The molecules of the
CC invention demonstrate cytosstatic activity whilst the method may useful
CC for detecting and differentiating between colon cell proliferative
CC disorders, including cancers such as colon adenoma and colon carcinoma.
CC The pNA (peptide nucleic acid)-oligomers are useful as probes for
CC determining cytosine methylation state or single nucleotide
CC polymorphisms. The current sequence is that of the pretreated genomic DNA
CC region of the invention. This sequence is not shown within the
CC specification but is taken from Wipoweb.

XX Sequence 11222 BP; 2767 A; 0 C; 2498 G; 5957 T; 0 U; 0 Other;

Query Match 7.3%; Score 105; DB 10; Length 11222;
Best Local Similarity 49.9%; Pred. No. 4.8e-06;

Matches 345; Conservative 0; Mismatches 340; Indels 6; Gaps 3;

QY 412 ATGGTACCTATAAAAAATTATTATTTTAAATAAATTTGATGACCAAGTGGTTGGAGAGA 471
DB 7223 ATGGTGTGTTGCTTAAGTTTGAATTGAATTAATGTTTATTGAGTAAATGGGTGGATGGT 7282

QY 472 GAGGTCTACCGATTGGTCAAGTGGCCCAATTTTATTATTTTACCTCGCTAGATTGCT 531
DB 7283 TTTGTAAGAAGAGAGAGTGGAGGATGATTAGTTTATTTTATTTATTTATTTATTTT 7342

QY 532 AATATCTATTGTCATTATCTCATTTTATTTTATTTTAAATTATTTTATTTTATTTTATTTT 589
DB 7343 AATTATTTAGGTATTTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7402

QY 590 TAAAAATTCCTAATCTTACTTTTTTTTAAAAAGAAATTTATTTAAATTTATTTATTTATTT 649
DB 7403 TATTTATTTATTTGTTATTTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7462

QY 650 TAGATAAAAATCTAATCTTACTTTTTTTTAAAAAGAAATTTCAATTCGGTTTTTCT 709
DB 7463 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7522

QY 710 TAATTTAGTTTTAATCTATCTAATTTATATAAAATTTCTGATCGATTAGTGTGCTCAA 769
DB 7523 TGAATTTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7582

QY 770 AGTCAAGTCAATGAATTTTGGAGAAAAATAAAAAATTAACACATTTTTTCGATTAA 829
DB 7583 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7642

QY 830 TTTTATTAATATAATAATAACACATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 889
DB 7643 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7702

QY 890 AATTAA - -ATTTCAGCACACAATTTACACTCTCATCATTAATAATTTTAAATCTTATTTACCAT 947
DB 7703 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7762

QY 948 AATTAAAAATTTGAGGCAATTTATTTTAAATCTCACCCCTCCCAATTAATGATATTTATTA 1007
DB 7763 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7822

QY 1008 TTTTGTTCGA - -TACTTCTATTTTCTCTCAATTAATTAATTAACCAATTTTGAAC 1065
DB 7823 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7882

QY 1066 TGTATTAATTTCTTAATCTTATTTACTATTGT 1096
DB 7883 TATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7913

AAF22281/c

ID AAF22281 standard; DNA; 59590 BP.

XX AAF22281;

AC AAF22281;

XX 20-MAR-2001 (first entry)

XX BAC containing repeats from centromeres 1-4 #4.

XX Centromere; microsome; vector; ds.

XX Arabidopsis thaliana.

XX WO200055325-A2.

XX 21-SEP-2000.

XX 17-MAR-2000; 2000WO-US007392.

XX 18-MAR-1999; 99US-0125219P.

PR 01-APR-1999; 99US-0127409P.

PR 18-MAY-1999; 99US-0134770P.

PR 13-SEP-1999; 99US-0153584P.

PR 17-SEP-1999; 99US-0154603P.

PR 16-DEC-1999; 99US-0172493P.

XX (UYCH-) UNIV CHICAGO.

XX Preuss D, Copenhaver G, Keith K;

PI WPI; 2000-587529/55.

XX Recombinant DNA construct comprising a plant centromere, useful for
PT producing stably inherited microsome which can serve as vectors for the
PT construction of transgenic plant and animal cells.

XX Claim 102; Page 351-364; 1449pp; English.

XX The present invention relates to a recombinant DNA construct of a plant
CC (Arabidopsis thaliana) centromere. The constructs are useful for
CC producing stably inherited microsome which can serve as vectors for the
CC construction of transgenic plant and animal cells expressing selected
CC proteins such as hormones, enzymes, interleukins, clotting factors,
CC cytokines, antibodies, and growth factors

XX Sequence 59590 BP; 17614 A; 12031 C; 13575 G; 16118 T; 0 U; 252 Other;

Query Match 7.3%; Score 105; DB 3; Length 59590;

Best Local Similarity 57.4%; Pred. No. 4.1e-06;

Matches 189; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 59 AITTCGATTAAGACTTCTTATCTATCAGAAGCGCTCGAGAGGATCCCAATTAGTCTAAA 118

DB 20299 ATCCGATTAAGACTTCTTCTGCGTTCGAGATACCTTTAACTTAACCTTAAGAAAGTTCAT 20240

QY 119 AITTCATCTCAGTCTCGGAAACCAACTCAGGACCCAAACCCGTCGCTCACCACCACTCAGT 178

DB 20239 CTTTTCATTCGTTCTCGTCCACAATCCCTGACCGAATAGCATTCATTTCGCCACGGGCA 20180

QY 179 CTAATATAACAGAGTATGACCTTATGACATATAGAGCTCGTAAGTGCCATCTAGAT 238

DB 20179 CCAACATAAGGGCATCCGGCATGCCCTTCTATACAGTGCCTCATAGGCGACATCACAAT 20120

QY 239 GCCAGATTCGAACTGTTATTTAGGGAACCTCACTAAACGGTAAAAAATCCCTCTCAACT 298

DB 20119 ACTTCTTGTAAGTACTGTTATTTATGATGCAAACTCTATCAAGCTCAATGCTTCTCAGATT 20060

QY 299 ACCTTAGTAATAATCACATAGCTCCAAATCGTATCTCTAGTATATGAATACCTTCTC 358

DB 20059 TCCACCCCAATCTAAACACATGCCCTTTTGATATCTTCTAGAGTCCGAATGTTGGCTC 20000

QY 359 AATTGACCATCGTCTGAGGATGGAATG 387

Db 12973 GCAAACTCTACCAGACTCAAGTGATCCGCCAATGACCTCCCAATCCCAAGACACACATC 13032
QY 324 CAAATCGTATCCCTCTAGTATATGAATCACTCTCAAAATGACCATCGGCTGAGGATCG 383
Db 13033 CGCAACATATCCCTCAAGTGT-TGAATAGTCTCTCAGATTCGCCATCTGTATGAGGATGA 13091
QY 384 AATGCAG 390
Db 13092 TAAGCCG 13098

RESULT 15
ID ACF62794
AC ACF62794 standard; DNA; 8222 BP.
AC ACF62794;
XX 09-OCT-2003 (first entry)
XX Colon cancer analysis related genomic DNA SEQ ID NO:43.
DE Human; colon cancer; oestrogen receptor; myoglobin; p21; p27; p16; p53;
KW progesterone receptor; pcna; cea; cdc2; c-erbB2; methylation; CpG;
KW characterisation; classification; diagnosis; differentiation;
KW colon cell proliferative disorder; gene; ds.
XX Homo sapiens.
OS Synthetic.
XX WO2003014388-A2.
XX 20-FEB-2003.
XX 09-AUG-2002; 2002WO-EP008939.
XX 09-AUG-2001; 2001DE-01039283.
XX (BPIG-) EPIGENOMICS AG.
XX Distler J, Model F, Taubert H;
XX WPI; 2003-256600/25.
XX Determining methylation status of CpG dinucleotides using modified
PT genomic sequences, oligonucleotides and/or PNA-oligomers, useful in the
PT characterization, grading, staging and/or diagnosis of colon cancer.
XX Claim 22; Page 81-83; 219pp; English.

XX The present invention describes a method for determining the methylation
CC status of CpG dinucleotides within the genes for oestrogen receptor, p21,
CC p27, p16, progesterone receptor, myoglobin, pcna, cdc2, c-erbB2, p53
CC and/or cea, which comprises contacting the target nucleic acid with a
CC reagent that distinguishes between methylated and non-methylated CpG
CC dinucleotides, and determining from the methylation status of the CpG
CC positions the presence of a colon cancer. A set of oligomers or peptide
CC nucleic acid (PNA)-oligomers can be used as probes for determining the
CC cytosine methylation state and/or single nucleotide polymorphisms (SNP)
CC of a corresponding genomic DNA by analysis of a chemically pretreated
CC genomic DNA. The pretreated genomic DNA is useful for the determination
CC of the methylation status of a corresponding genomic DNA and/or detection
CC of SNPs. The methods and pretreated genomic DNA are also useful for the
CC characterisation, classification, diagnosis and differentiation of colon
CC cell proliferative disorders. ACF62752 to ACF63278 represent sequences
CC used in the exemplification of the present invention

XX Sequence 8222 BP; 2010 A; 116 C; 1769 G; 4327 T; 0 U; 0 Other;
XX Query Match 7.2%; Score 103.4; DB 8; Length 8222;
XX Best Local Similarity 49.8%; Pred. No. 8.3e-06;
XX Matches 344; Conservative 0; Mismatches 341; Indels 6; Gaps 3;
QY 412 ATGGTACCTATAAAAAATTATTATTTTAAAAAAATGTGACCAAGTGGTGGAGAGA 471

Db 7223 ATGGTGTGTTTGTGTTTAAAGTTTGAATTGAATAAATGTTTATTGAGTAAATGGGTGGATGTT 7282
QY 472 GAGGCTACCGATTGGTCAAGTGGCACCACAAATTTTATTATTACCTCCCTGAGATTGCT 531
Db 7283 TTTGTAAGAAAGAGAGAGTGGAGATGATTAGTTTATTATTATTATTATTATTATTATT 7342
QY 532 AAATACTATTGCAATTTATCTCATTTTCATTTTCAATTTTATTATTATTATTATTATT 589
Db 7343 AATTATTATTAGGTATTAAATTTATTATTATTATTATTATTATTATTATTATTATTATT 7402
QY 590 TAAAAATCTCAATACCTTACTTTTAAAAAGAAATTTTAAATTTTAAATTTTAAATTTT 649
Db 7403 TATTATTATTATTGTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 7462
QY 650 TAGATAAAAATCTAATACTTTTACTTTTAAAAAGAAATTTCAATTTGCGTCTTTTCT 709
Db 7463 TTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7522
QY 710 TAAATTTAGTTTAAATTTCTATCTAATTTATAAAAAATTTCTGATCGGATAGTGGTCAA 769
Db 7523 CGATATTTTTAAATTTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7582
QY 770 AGTCAAGTCACATGAATTTTGTGGAGAAAAATAAAAATTAACACATTTTTCGATTAA 829
Db 7583 TTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7642
QY 830 TTTATTATTATTATAATAATAAACAACATTTTATTATTATTATTATTATTATTATTATTATT 889
Db 7643 TTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7702
QY 890 AATTAAAA--ATTTTCAGCACACAAATTTACACTCTCATCATTAATAATTTAACTTTATACCAT 947
Db 7703 TATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7762
QY 948 AATTAAAAATGTGAGACAAATTTTAAATCTCACCCCTCATTAATGCAATTAATTAATAA 1007
Db 7763 TTTTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7822
QY 1008 TTTTGTGTCGA--TACTTCTTATTTCACCTCCTAACACATTAATCAATTAACCAATTTTGAAC 1065
Db 7823 TTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7882
QY 1066 TGTATAAATTTCTTAACCTTTATTCACCTATTGT 1096
Db 7883 TATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 7913

Search completed: August 13, 2005, 08:07:31
Job time : 1238 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 12, 2005, 18:51:23 ; Search time 10337 seconds
(without alignments)
6717.261 Million cell updates/sec

Title: US-10-089-543-2

Perfect score: 1433

Sequence: 1 actataggacgcgtgtgtc.....tgagagaatccttcacatc 1433

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1433	100.0	1433	6	AX370650 Sequence
2	1393.8	97.3	5765	8	AF487511 Gossypium
C 3	266	18.6	572	8	AF025353 Gossypium
C 4	184	12.8	300	8	AF060583 Gossypium
C 5	179.4	12.5	5998	8	AY181254 Gossypium
C 6	179	12.5	913	8	AF276939 Gossypium
C 7	156.4	10.9	927	8	AF277099 Gossypium
8	150.6	10.5	111063	2	AC146584 Medicago
9	140.2	9.8	121257	2	AC146307 Medicago
10	139.8	9.8	119270	2	AY508219 Medicago
11	139	9.7	96930	8	AC149484 Populus b
C 12	137.6	9.6	78069	2	AC150760 Medicago
C 13	137.4	9.6	141416	8	AC149480 Populus b
C 14	137.2	9.6	143392	8	AC149299 Populus b
C 15	136.4	9.5	107287	2	AC146794 Medicago
C 16	136.4	9.5	181249	2	AC148479 Zea mays
17	135.6	9.5	65127	2	AC148403 Medicago
C 18	135.6	9.5	123241	2	AY379775 Medicago
19	135.4	9.4	106114	2	AC146565 Medicago

C 20	135.4	9.4	122167	8	AC146806	AC146806 Medicago
C 21	135	9.4	105035	2	AC145452	AC145452 Zea mays
C 22	134	9.4	170017	2	AC148158	AC148158 Zea mays
C 23	133.8	9.3	127841	2	AC147714	AC147714 Medicago
C 24	133.6	9.3	100595	8	AP004545	AP004545 Lotus cor
C 25	133.4	9.3	116840	2	AC141862	AC141862 Medicago
C 26	133.4	9.3	125529	2	AC121242	AC121242 Medicago
27	132.4	9.2	116004	8	AC149208	AC149208 Medicago
28	132.4	9.2	135726	8	AC147498	AC147498 Medicago
29	132.2	9.2	125478	8	AC147430	AC147430 Medicago
C 30	132	9.2	125522	2	AC148718	AC148718 Medicago
C 31	132	9.2	198102	2	AC145389	AC145389 Zea mays
C 32	131.8	9.2	86095	8	AC136450	AC136450 Medicago
C 33	131.2	9.2	113049	8	AC146649	AC146649 Medicago
C 34	131	9.1	185534	2	AC145481	AC145481 Zea mays
35	130.6	9.1	78128	8	AC144805	AC144805 Medicago
C 36	130.6	9.1	119996	2	AC137666	AC137666 Medicago
C 37	130.4	9.1	171261	2	AC146349	AC146349 Zea mays
C 38	130.4	9.1	178675	2	AC145262	AC145262 Zea mays
39	130.2	9.1	44109	2	AC146805	AC146805 Medicago
40	130.2	9.1	110625	2	AC138199	AC138199 Medicago
C 41	130.2	9.1	124889	8	AC126008	AC126008 Medicago
C 42	130.2	9.1	133526	2	AC135315	AC135315 Medicago
C 43	130.2	9.1	137686	2	AC130806	AC130806 Medicago
44	130	9.1	206462	2	AC148173	AC148173 Zea mays
45	129.6	9.0	132211	8	AC135396	AC135396 Medicago

ALIGNMENTS

RESULT 1
AX370650 LOCUS AX370650 1433 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 2 from Patent WO0210377.
ACCESSION AX370650
VERSION AX370650.1 GI:19168815
KEYWORDS
SOURCE Arabidopsis sp.

ORGANISM Arabidopsis sp.
Eukaryote, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, rosids, eurosids II, Brassicales, Brassicaceae, Arabidopsids.

REFERENCE 1
AUTHORS Cai, L., Li, X., Cheng, N. and Liu, J. W.

TITLE Isolation and characterization of a fiber-specific g(b)-tubulin

JOURNAL Promoter from cotton

PATENT: WO 0210377-A 2 07-FEB-2002;

Institute of Molecular Agrobiolgy (SG)

FEATURES
Location/Qualifiers

source 1..1433

/organism="Arabidopsis sp."

/mol_type="unassigned DNA"

/db_xref="taxon:29726"

ORIGIN

Query Match 100.0%; Score 1433; DB 6; Length 1433;

Best Local Similarity 100.0%; Pred. No. 7.4e-189;

Matches 1433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACTATAGGCACGCGTGTGTGACGCGCGCGGCTGCTGATATCTATGATTTTCAGAT 60

|||||

Db 1 ACTATAGGCACGCGTGTGTGACGCGCGCGGCTGCTGATATCTATGATTTTCAGAT 60

|||||

Qy 61 TTGCATTAAGACTTCTATCTATCAGAACGCGCTCAGAGGATCCCAATTAAGTCTAAAT 120

|||||

Db 61 TTGCATTAAGACTTCTATCTATCAGAACGCGCTCAGAGGATCCCAATTAAGTCTAAAT 120

|||||

Qy 121 TATCTTCAGTCTCGGAACCAACTCAGACCCCAAAACCCGCTGCTCACCACCTCAGTCT 180

|||||

Db 121 TATCTTCAGTCTCGGAACCAACTCAGACCCCAAAACCCGCTGCTCACCACCTCAGTCT 180

|||||

Qy 181 AATATAACAGATATGACACTTATGACCATATAGAGCCTCGTAAGGTGCCATCTAGATGC 240

|||||

181 AATATAACAGATGATGACATTATGACCATATAGAGCCTCGTAGGTCGCATCTAGATGC 240
241 CAGATTGGAACCTGTTATTGTTAGCGGAACTCACTAACGGTAAATAATCCTCTCAACTAC 300
241 CAGATTGGAACCTGTTATTGTTAGCGGAACTCACTAACGGTAAATAATCCTCTCAACTAC 300
301 CTTAGTAAATAACACATAGCTCAAAATCGTATCTCTAGTATATGAATCACCTTCTCAA 360
301 CTTAGTAAATAACACATAGCTCAAAATCGTATCTCTAGTATATGAATCACCTTCTCAA 360
361 ATTGACCATCGGTCGAGGAGGAAATGACACACCGGTGCCACCGATTAATATGTTACCT 420
361 ATTGACCATCGGTCGAGGAGGAAATGACACACCGGTGCCACCGATTAATATGTTACCT 420
421 ATAAAAAATTAATATTTTAAAAAATTTGATGACACAGGTTGGAGAGAGGTTCTAC 480
421 ATAAAAAATTAATATTTTAAAAAATTTGATGACACAGGTTGGAGAGAGGTTCTAC 480
481 CGATTGCTCAAGTGGACCAATTTTATTTTATTTTACCTCCTGCTAGATTCTGAATACTAT 540
481 CGATTGCTCAAGTGGACCAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 540
541 TGCATTATCTCAATTCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 600
541 TGCATTATCTCAATTCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 600
601 ATACTTTACTTTTAAAAAGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 660
601 ATACTTTACTTTTAAAAAGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 660
661 TCTAATCTTTACTTTTAAAAAGAAATTTCAATTTGCTGTTTCTTAAATTTAGTTT 720
661 TCTAATCTTTACTTTTAAAAAGAAATTTCAATTTGCTGTTTCTTAAATTTAGTTT 720
721 TAATCTTACTAATTAATAAAATCTGATCGGAAATGAGTGGTGCACAAAGTCAAGTCAC 780
721 TAATCTTACTAATTAATAAAATCTGATCGGAAATGAGTGGTGCACAAAGTCAAGTCAC 780
781 ATGAATTTTCTGAGAGAAAAATAAAATTAACACATTTTTCGATTAATTTATTTATTA 840
781 ATGAATTTTCTGAGAGAAAAATAAAATTAACACATTTTTCGATTAATTTATTTATTA 840
841 TATAAATAATAACACATTTTATTTAATGTTCTCAATAATTTTAAATTTTAAATTT 900
841 TATAAATAATAACACATTTTATTTAATGTTCTCAATAATTTTAAATTTTAAATTT 900
901 CAGCACAAATTAACATCTCATCATTAATTAATCTTATTAATCAATTAATTAATTAATTTGTG 960
901 CAGCACAAATTAACATCTCATCATTAATTAATCTTATTAATCAATTAATTAATTAATTTGTG 960
961 AGGCAATTAATTTTAAATCTCACTCCCTCCATTAATGATATTAATTTTGTTCGATA 1020
961 AGGCAATTAATTTTAAATCTCACTCCCTCCATTAATGATATTAATTTTGTTCGATA 1020
1021 CTTCTTATTTCACTCTCAATTAATTAATCAATTAATTTTGAATCTGTTAATTTCTTA 1080
1021 CTTCTTATTTCACTCTCAATTAATTAATCAATTTTGAATCTGTTAATTTCTTA 1080
1081 ACTTATTTCAATTTGTTGCTCTGGTGCATCTGGAAGGCGACCGTCCAGGCTGTCCAAC 1140
1081 ACTTATTTCAATTTGTTGCTCTGGTGCATCTGGAAGGCGACCGTCCAGGCTGTCCAAC 1140
1141 CACACTTTGCCAGCTGATCAATTTCCAGTAATCTATGTTTACAGTACTTAAAGCAATTTCC 1200
1141 CACACTTTGCCAGCTGATCAATTTCCAGTAATCTATGTTTACAGTACTTAAAGCAATTTCC 1200
1201 AATTTCAAAATTTCAATTTCCAGGAAAAACGAAACGTCGTTACTAACCGACCTTAAACC 1260
1201 AATTTCAAAATTTCAATTTCCAGGAAAAACGAAACGTCGTTACTAACCGACCTTAAACC 1260
1261 CAGCTCAACCTGCGTCAATTAACGGAATCTTTTAACTCTCTATATTAACCCAAACCA 1320

1261 CAGCTCAACCTGCGTCAATTAACGGAATCTTTTAACTCTCTATATTAACCCAAACCA 1320
1321 CTTCTCATCACTATTTCCCATATAAAGAAATTTCCGGAATTTCTTATTTCTTTATATTTT 1380
1321 CTTCTCATCACTATTTCCCATATAAAGAAATTTCCGGAATTTCTTATTTCTTTATATTTT 1380
1381 CTTCTCAATTTCCCTGCTCACTTTCCGAGAAAAATGAGAAAAATCCTTTCATC 1433
1381 CTTCTCAATTTCCCTGCTCACTTTCCGAGAAAAATGAGAAAAATCCTTTCATC 1433

RESULT 2
AF487511 5765 bp DNA linear PLN 17-OCT-2002
LOCUS Gossypium hirsutum beta-tubulin (TUB1) gene, complete cds.
DEFINITION AF487511
ACCESSION AF487511
VERSION AF487511.1 GI:19569608
KEYWORDS Gossypium hirsutum (upland cotton)
SOURCE Gossypium hirsutum
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 5765)
AUTHORS Li, X.B., Cai, L., Cheng, N.H. and Liu, J.W.
TITLE Molecular Characterization of the Cotton GhTUB1 Gene That Is Preferentially Expressed in Fiber
JOURNAL Plant Physiol. 130 (2), 666-674 (2002)
MEDLINE 22263995
PUBMED 12376634
REFERENCE 2 (bases 1 to 5765)
AUTHORS Li, X.B., Cai, L. and Liu, J.W.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2002) Laboratory of Plant Reproduction and Cotton Biotechnology, Institute of Molecular Agrobiolgy, 1 Research Link, NUS, Singapore 117604, Republic of Singapore

FEATURES
source
1..5765
/organism="Gossypium hirsutum"
/mol_type="genomic DNA"
/db_xref="taxon:3635"
<1377..>3349
/gene="TUB1"
join(<1377..1771,1876..2144,2676..>3349)
/gene="TUB1"
/product="beta-tubulin"
join(1377..1771,1876..2144,2676..3349)
/gene="TUB1"
/note="microtubule protein subunit"
/codon_start=1
/product="beta-tubulin"
/protein_id="AAL92118.1"
/db_xref="GI:19569609"
/translation="MREILHIOGGCGNOIGAKFWFVCAEHGIDSTGRYGGDELQ
ERINLYNEASCRFPVRAVLMLEPGTWDMSVRSQYQIIFPDNFRVGGSGAGNWA
KHYTGAEILIDVLDVVRKAENCDLQGFQVCHSLGGTSGMGILLISKRBEYP
DRMLTFVFPSPKSDTVVEPYNATLSVHQLVENADECMVLNDALYDLCFTKLUT
TPSGGLNHLISATMSGVTCCLFPPQLNSDLKLAIVNLIFFRLHFFMVGFAPLTSR
GSQOYRALTVPELTQQMDAKNNMCAADPRHRYLTASAVRKGKMTKEVDEQMINVQ
KNSSYFVSWIPNNVKSVCDDIPPIGLKMASTFIGNSTSIQEMPRRVSEQFTAMFRRK
AFLHWTTGEGMDEMEFTEAESNNDLVSEYQQYQDATADDEEYEEBEYEA"

ORIGIN
Query Match 97.3%; Score 1393.8; DB 8; Length 5765;
Best Local Similarity 99.9%; Pred. No. 1.4e-183;
Matches 1395; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 37 GCTTGATATCTATGATTTTCAGATTTGCAATAAGACTTCTATCTATCAGAAGACGCTTGA 96
Dd 1 GCTTGATATCTATGATTTTCAGATTTGCAATAAGACTTCTATCTATCAGAAGACGCTTGA 60
Qy 97 GAGGATCCCAAAATAGTCTAAATTTATCTTCAGTCTCGGAAACCAACTCAGGACCCAA 156

Db 61 GAGGATCCCAATAGTCTAAAAATTATCTTCAGTCTCGGAAACCAACTCAGGACCCAAAA 120
Qy 157 CCCGTCGCTCACCCAACTCAGTCTAAATATAACAGAGTATGACACTTATGACCATATAGAG 216
Db 121 CCCGTCGCTCACCCAACTCAGTCTAAATATAACAGAGTATGACACTTATGACCATATAGAG 180
Qy 217 CCTCGTAAGGTGCCATCTAGATGCCAGATCGGAACCTGTTATGTAGGCGAACTCAACTA 276
Db 181 CCTCGTAAGGTGCCATCTAGATGCCAGATCGGAACCTGTTATGTAGGCGAACTCAACTA 240
Qy 277 ACGGTAAAAATCTCTCAACTACCTTAGTATTAATAACACATAGCTCCAAATCGTATCCT 336
Db 241 ACGGTAAAAATCTCTCAACTACCTTAGTATTAATAACACATAGCTCCAAATCGTATCCT 300
Qy 337 CTAGTATATGAATCAACCTTCTCAAAATGACCATCGGTCTGAGGATGGAATGCAGACCGGT 396
Db 301 CTAGTATATGAATCAACCTTCTCAAAATGACCATCGGTCTGAGGATGGAATGCAGACCGGT 360
Qy 397 GCCACCGATTTACTAATGGTACCTATAAAAAATTTATTTTAAAAAATTTGATGTGAC 456
Db 361 GCCACCGATTTACTAATGGTACCTATAAAAAATTTATTTTAAAAAATTTGATGTGAC 420
Qy 457 CAGTGGTTGGAGAGAGGCTCTACCGATTTGTCAGTGGGACCAATTTTATTTTACCT 516
Db 421 CAGTGGTTGGAGAGAGGCTCTACCGATTTGTCAGTGGGACCAATTTTATTTTACCT 480
Qy 517 CCTGCTAGATTCGTAATACCTATTGCAATTTATCTCATTTTCATTTTATTTTAAATTT 576
Db 481 CCTGCTAGATTCGTAATACCTATTGCAATTTATCTCATTTTCATTTTATTTTAAATTT 540
Qy 577 TATATTATTTGGATAAAAAATTTCTAATCTTTACTTTTTTTTTTAAAAAAGAAATTTTAAAT 636
Db 541 TATATTATTTGGATAAAAAATTTCTAATCTTTACTTTTTTTTTTAAAAAAGAAATTTTAAAT 600
Qy 637 ATTTTATATTTAGATAAAAAATTTCTAATCTTTACTTTTTTTTTTAAAAAAGAAATTTCAA 696
Db 601 ATTTTATATTTAGATAAAAAATTTCTAATCTTTACTTTTTTTTTTAAAAAAGAAATTTCAA 660
Qy 697 TTGGTGTCTTCTAATTTAGTCTTAAATTTAATTTCTATCTAATTAATAAAATTTCTGATG 756
Db 661 TTGGTGTCTTCTAATTTAGTCTTAAATTTAATTTCTATCTAATTAATAAAATTTCTGATG 720
Qy 757 AGTGTGTGTCAAAGTCAAGTCAATGAATTTTGTGGAGAAAAATTAATAAATTTAAACAC 816
Db 721 AGTGTGTGTCAAAGTCAAGTCAATGAATTTTGTGGAGAAAAATTAATAAATTTAAACAC 780
Qy 817 ATTTTTCGATTAATTTAT 876
Db 781 ATTTTTCGATTAATTTAT 840
Qy 877 AATAATATTTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 936
Db 841 AATAATATTTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 900
Qy 937 CTATTATACCAATTTAAATTTGTGAGCAATTTATTTTAAATCTCACCTCCCAATTAATG 996
Db 901 CTATTATACCAATTTAAATTTGTGAGCAATTTATTTTAAATCTCACCTCCCAATTAATG 960
Qy 997 CATATTATTAATTTTGTTCGATCTTCTTATTTTCACTCTCTTAAATTAATTAATTAACCCA 1056
Db 961 CATATTATTAATTTTGTTCGATCTTCTTATTTTCACTCTCTTAAATTAATTAATTAACCCA 1020
Qy 1057 ATTTTGAACCTGTTAATTTTCTTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1116
Db 1021 ATTTTGAACCTGTTAATTTTCTTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1080
Qy 1117 AGGCCACCGTCTCAGGCTGTCCAAACCACTTTTGCACGTCTCAATTTCCAGTAACTACAT 1176
Db 1081 AGGCCACCGTCTCAGGCTGTCCAAACCACTTTTGCACGTCTCAATTTCCAGTAACTACAT 1140
Qy 1177 TGTTACAGTTACTAAGCAAAATCCCAATTTCAAAATTTCAAAATTTCCAGGAAAAACGAAACG 1236
Db 1141 TGTTACAGTTACTAAGCAAAATCCCAATTTCAAAATTTCAAAATTTCCAGGAAAAACGAAACG 1200

Qy 1237 TCCGTTACTAACCGACCTAAACCCAGCTCAACCTGCCGTCAATTAACGGAATCTTTTA 1296
Db 1201 TCCGTTACTAACCGACCTAAACCCAGCTCAACCTGCCGTCAATTAACGGAATCTTTTA 1260
Qy 1297 ACTCCTCTATATAACCCAAACCACTCTCATCACCATTTCCTCCATTAAGAAATTTCCGG 1356
Db 1261 ACTCCTCTATATAACCCAAACCACTCTCATCACCATTTCCTCCATTAAGAAATTTCCGG 1320
Qy 1357 AATTTCTTATTTCCCTTTTATATTTTCCCTCTCCAAATTTCCCGTCACTTTCCCGAGAAAATGA 1416
Db 1321 AATTTCTTATTTCCCTTTTATATTTTCCCTCTCCAAATTTTCCTGTCACCTTTCCCGAGAAAATGA 1380
Qy 1417 GAGAAATCTTTTCATC 1433
Db 1381 GAGAAATCTTTTCATC 1397

RESULT 3
AF025353/c 572 bp DNA linear PLN 08-JAN-1999
LOCUS Gossypium hirsutum tandem repeat B77 sequence.
DEFINITION AF025353
ACCESSION AF025353
VERSION AF025353.1 GI:4128202

KEYWORDS
SOURCE Gossypium hirsutum (upland cotton)
ORGANISM Gossypium hirsutum

REFERENCE 1 (bases 1 to 572)
AUTHORS Zhao X., Ji Y., Ding X., Stelly, D.M. and Paterson, A.H.
TITLE Macromolecular organization and genetic mapping of a rapidly evolving chromosome-specific tandem repeat family (B77) in cotton (Gossypium)

JOURNAL Plant Mol. Biol. 38 (6), 1031-1042 (1998)
MEDLINE 98084754
PUBMED 9869409

REFERENCE 2 (bases 1 to 572)
AUTHORS Zhao X., Ji Y., Ding X., Stelly, D.M. and Paterson, A.H.
TITLE Direct Submission

JOURNAL Submitted (12-SEP-1997) Dermatology, University of Michigan, 1500 E. Medical Center Dr., 3420 CCGC, Box 0932, Ann Arbor, MI 48109-0932, USA

FEATURES
Location/Qualifiers
source 1..572
/organism="Gossypium hirsutum"
/mol_type="genomic DNA"
/db_xref="taxon:3635"

repeat_region 1..572
/rpt_family="B77"
/rpt_type=tandem

ORIGIN
Query Match 18.6%; Score 266; DB 8; Length 572;
Best Local Similarity 91.6%; Pred. No. 1.4e-27;
Matches 326; Conservative 0; Mismatches 25; Indels 5; Gaps 4;

Qy 35 GTGCTTGATATCTATGATTTTTCAGATTTGCATAGACTTCTATCTATCAGAGAGCGCTG 94
Db 421 GTACTGAAATCTCTGA-TTTCAGATATGCATAGACTTCTATCTATCAGAGAGCGCTT 363

Qy 95 CAGAGGATCCCAATTTAGTCTTAAATTTATCTTCAGTCTCGGAAACCAACTCAGGACCCAA 154
Db 362 CAC--GATCCCAATTTAGTCTTAAATTTATCTTCAGTCTCGGAAACCAACTCAGGAC 306

Qy 155 AACCCGTCGCTCACCCAACTCAGTCTTAATATAACAGAGTATGACATTTATGACCATATAG 214
Db 305 AACCCGTCGCTCACCCAACTCAGTCTTAATATAACAGAGTATGACATTTATGACCATATAG 247

Qy 215 AGCCTCGTAAGGTGCCATCTAGATGCGAGATGGAACCTGTTATTTAGGGCAACTCAAC 274
Db 246 AACCTCGTAAGGTGCCATCTAGATGCGAGATGGAACCTGTTATTTAGGGCAACTCAAC 187

shape differences in cotton (*Gossypium hirsutum* L.)
Unpublished
REFERENCE 2 (bases 1 to 913)
AUTHORS Ulloa,M. and Meredith,W.R. Jr.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-2000) C. G. & P. Research Unit, USDA-ARS, 141 Experiment Station Rd. Box 345, Stoneville, MS 38776, USA

FEATURES
source 1. .913
/organism="Gossypium hirsutum"
/mol_type="genomic DNA"
/db_xref="taxon:3635"
/clones="Cott-2"
repeat_region <1. .913
/note="similar to retrotransposon del 1-46 from *Lilium henryi*"
misc_feature <1. .635
/note="nonfunctional integrase due to mutation; putative zinc finger"

ORIGIN
Query Match 12.5%; Score 179; DB 8; Length 913;
Best Local Similarity 70.5%; Pred. No. 1.3e-15;
Matches 239; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Qy 42 ATATCTATGATTTTCAGATTTCGATAGACTTCTATCTATCAGAGAGCGCTGCAGAGGA 101
Db 340 ATATTCTTTCTTTTCAAGTCTGCATAGACTTTTGTATTATCGGAGGTGCTTTTCAGACGA 281

Qy 102 TCCCAAAATAGTCTAAATATATCTTCAGTCTCGGAAACCAACTCAGGACCCCAAAACCCGT 161
Db 280 TCAGATACATTTTCATTTGTTCTCAGTCTCGGAACCACTTTAGGACCCCAATACTCAA 221

Qy 162 CGCTCACCACACTCAGTCTAATAAACAAGATGACACTTATGACCAATATAGAGCTCG 221
Db 220 CTTTCACCTTAATTCATCAACATCAACAGAGTAGCAGACTACCAACATATAAGGCGCTCG 161

Qy 222 TAAGGTGCATCTAGATGCCAGATTGGAACTGTTATTGTAGGCGAACTCAACTAAGCGT 281
Db 160 TAGGGTGCCAAATGAATACTAGATTAGAACTGTTATTGTATGGAACTCAGCAGTGCG 101

Qy 282 AAAAAATCCTCTCAACTACCTTAGTAATAAATACATAGCTCCAAATCGTATCTCTAGT 341
Db 100 AGTACTCGTCCCACTACCTTGAAGTCTATACACAACTTCGGAGCATATCTCTAAG 41

Qy 342 ATATGAATCACCTTCTCAAAATGACCATCGGTCTGAGGA 380
Db 40 ATTTGAATAACTCGCTCGCACTATCCATCAGTCTGGGA 2

RESULT 7
AF277099/c 927 bp DNA linear PLN 02-JUL-2003
LOCUS *Gossypium hirsutum* clone Cott-3 retrotransposon Ty3-Gypsy-like reverse transcriptase-like sequence.
DEFINITION
ACCESSION AF277099
VERSION AF277099.1 GI:32400675
KEYWORDS
SOURCE
ORGANISM *Gossypium hirsutum* (upland cotton)
Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; *Gossypium*.
1 (bases 1 to 927)
Ulloa,M. and Meredith,W.R. Jr.
TITLE A family of retrotransposon-like sequences associated with leaf shape differences in cotton (*Gossypium hirsutum* L.)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 927)
AUTHORS Ulloa,M. and Meredith,W.R. Jr.
TITLE Direct Submission
JOURNAL Submitted (13-JUN-2000) C.G. & P. Research Unit, USDA-ARS, 141 Experiment Station Rd. Box 345, Stoneville, MS 38776, USA

FEATURES
Location/Qualifiers
source 1. .927
/organism="Gossypium hirsutum"
/mol_type="genomic DNA"
/db_xref="taxon:3635"
/clones="Cott-3"
misc_feature 1. .927
/note="similar to reverse transcriptase sequence from *Arabidopsis thaliana* encoded by GenBank Accession Number AF077407"
repeat_region <1. .927
/transposon="Ty3-Gypsy-like retrotransposon"

ORIGIN
Query Match 10.9%; Score 156.4; DB 8; Length 927;
Best Local Similarity 66.3%; Pred. No. 1.8e-12;
Matches 240; Conservative 0; Mismatches 121; Indels 1; Gaps 1;

Qy 20 CGACGGCCGGGCTGGTGTGATATCTATGATTTTCAGATTTGCAATAAGACTTCTATCT 79
Db 363 CAAAGTCCCCCATAGAGATTCAATTCCTTACGTTCCAAATTCGATATGACTTCTGTTT 304

Qy 80 ATCAGAGAGCGCTGCAGAGGATCCCAAAATAGTCTAAATATTATCTTCAGTCTCGGAAC 139
Db 303 GTCAGATGCTTTCTTCAATTAGTTTCGAATCAGTTTACCTTCTCTTAGGTATCAGAAAC 244

Qy 140 CAATCTAGGACCCCAAAACCGCTCGCTCACCACTCAGTCTAATATAACAGAGTATGACA 199
Db 243 TAACTTGGGCCCGAGGACTCGCTCTCGCTAGCTAGTCAACAGAGTATGAGTATGACA 184

Qy 200 CTTATGACCATATAGAGCTCGTAAGGTGCCATCTAGATGCCAGATTTGGAACCTGTTATT 259
Db 183 CCTTTGACCATATAACGCTCGTACGCTGCAATTTGAATTAATAGTATGAGTATGTTATT 124

Qy 260 GTAGGCGAATCAACTAACGTTAAAAATCTCTCAACTACCTTAGTAAATAATACACATA 319
Db 123 ATATGCGAATCTGCTAGCGTCAAAATAGTTTTCCTCAACTGCTCTCAGAAATCTATTACACA 64

Qy 320 GCTCC-AAATCGTATCTCTAGTATATGAATCACTTCTCAAAATGACCATCGGTCTGAG 378
Db 63 ACCTTCAACTATATCATCGGTACTTGAATCACTCTTCGATGACCATCTGCTGCGG 4

Qy 379 GA 380
Db 3 GA 2

RESULT 8
AC146584 111063 bp DNA linear HTG 17-SEP-2004
LOCUS *Medicago truncatula* clone mth2-68k18, WORKING DRAFT SEQUENCE, 6 ordered pieces.
DEFINITION
ACCESSION AC146584
VERSION AC146584.28 GI:52219258
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE *Medicago truncatula* (barrel medic)
ORGANISM *Medicago truncatula*
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; *Trifolium*.
1 (bases 1 to 111063)
Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
TITLE *Medicago truncatula* BAC Clone mth2-68k18
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 111063)
AUTHORS Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-2003) Department of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

```
OK 73019, USA
3 (bases 1 to 111063)
Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
Direct Submission
Submitted (17-SEP-2004) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Sep 17, 2004 this sequence version replaced gi:52000565.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 55923: contig of 55923 bp in length
* 55924 56023: gap of unknown length
* 56024 87161: contig of 31138 bp in length
* 87162 87261: gap of unknown length
* 87262 92041: contig of 4780 bp in length
* 92042 92141: gap of unknown length
* 92142 105791: contig of 13650 bp in length
* 105792 105891: gap of unknown length
* 105892 108659: contig of 2768 bp in length
* 108660 108759: gap of unknown length
* 108760 111063: contig of 2304 bp in length.
Location/Qualifiers
1. 111063
/organism="Medicago truncatula"
/mol type="genomic DNA"
/db xref="taxon:3880"
/clone="mth2-68kl8"
/clone_lib="Medicago truncatula BAC library H2"

Query Match 10.5%; Score 150.6; DB 2; Length 111063;
Best Local Similarity 49.9%; Pred. No. 4.1e-12;
Matches 440; Conservative 0; Mismatches 429; Indels 12; Gaps 2;

Qy 27 CCGGGCTGGCTTGATCTATCTATGATTTTCAGATTTCGATTAAGACTCTCTATCTATCAAGAA 86
Db 41044 CTTTCCTGAAACTCAAGATCTTTTCTACGCTTATCATGATAAATCTTTTGTGCACTCTGC 41103

Qy 87 GAGCCTGCGAGAGATCCCAATTAGTCTAAATATCTTCAATCTTCAGTCTCGGAACCACTCA 146
Db 41104 GACGCTTTCATTTCTCTAGATCATCTGAACTCTCTCAGTAGTTTCTGCGAACAATCTCT 41163

Qy 147 GGACCCAAACCCGCTGCTCACCACCTCACTCTAATATACAGAGATGACACTTATGA 206
Db 41164 GGTCTTAGACCATCTTTACCTGACTCAACACGACACAGGAGTTCTGCATCTCGGA 41223

Qy 207 CCATATAGAGCCTCGTAGGTCGCATCTAGATGCGAGATTGGAACATGTTATTGTAGCGC 266
Db 41224 CCATATAAGGCTCGAAAGTGCCATCCAACTACTGGAATGATAGCTATTATTGTATGTG 41283

Qy 267 AACTCAACTACGCTAAATAATCTCTCACTACTTAGTAAATAAATCACTAGCTCCAA 326
Db 41284 AACTCGATCAACGGAAGATGACTATCCCAAGTTCCTCTGCTCAAGAACAAATCCTC 41343

Qy 327 ATCGTATCTCTAGTATATGATCACTCTCTCAAAATTCACCATCGGCTCTGAGGATGAAT 386
Db 41344 AACAAATCTCTAGGACTGAATGTCCTCTCCGACTGACCATCTATCTGTGGATGATAC 41403

Qy 387 GCAGACGGTGCCACCGATTACTTAATGGTACCTATATAAATAATATTATTTTAAAAAA 446
```

```
Db 41404 GCCGAACCTCAATCTCAACTTTCACATAATGTATCAAGTGTGAATTTGAAGTCTTTAAAAAAA 41463
Qy 447 TTGATGTGACCACTGGTGTGGAGAGAGAGGTCTACCGATTGGTCAAGTGGCACCACAAATTTT 506
Db 41464 TGTGAATTAAGATTATATATTGGATGAATAGTAGAATTTGAACA-----ACAAATAAGT 41518
Qy 507 TATTTTACCTCCGCGCTAGATTGCGTAATAACTATGATTCGATTTATCTCATTTCTATTTAT 566
Db 41519 GAGATGACCCAVATAATACTAAATTTCTTAAGATTTTAGGGTGACATGCGATATCAAACTCGAT 41578
Qy 567 TTAATTTATTTTATTTATTTTGGATAAAATTTCTAATACCTTTACTTTTTTTAAAAAGAAT 626
Db 41579 GTAGGATACCGGAAGAAGTTCCAAAATATATGATGAATTTTGTGTAGTTTTTATGATATCA 41638
Qy 627 TTAATTTAAATTTATTTATTTATTTAGATAAAATTTCTAATACCTTTACTTT-----TTT 679
Db 41639 TTTTAAATTTTTCATCTTAATTTTGTATTGATGATGTCATTTCACTTTTCATGCTATTT 41698
Qy 680 TTTAAAAAGAAATTTCAATTTGCGTTTTTCTTAAATTTAGTTTTTAATTTCTATATCAATATA 739
Db 41699 TTTTACTCAACCTCCATTTTTTACGCTTTTGTGAAATTTTAAATGACGTGACGTTT 41758
Qy 740 AAAATTTCTGATCGGATTAGTGTCTCAAGTCAAGTCACATGAATTTTGTGGAGAAA 799
Db 41759 TTAATTTGAGCCTTAATTTTATAAAATAATATCTCAAAATCTTCAATCAATCAACAA 41818
Qy 800 AAATAAAAAATTAAACACATTTTTCGATTAATTTATTTATATATATAATAATAATAACACAT 859
Db 41819 AAAAATCTACAAAAATATCACTCACTTCATTTCTTATATATCTCGGTTAAACACAA 41878
Qy 860 TTTTATTTAATGTTGTCAATAATATTTTAAATAAAATTT 900
Db 41879 TTTTCTTCTATCACTTCTCTCATGTTAAATTTGATTTT 41919

RESULT 9
AC146307 121257 bp DNA linear HTG 09-SEP-2004
LOCUS Medicago truncatula clone mth2-9p17, WORKING DRAFT SEQUENCE, 3
DEFINITION ordered pieces.
ACCESSION AC146307 GI:51948617
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT.
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 121257)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Medicago truncatula BAC Clone mth2-9p17
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 121257)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-2003) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 121257)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-2004) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Sep 9, 2004 this sequence version replaced gi:51699618.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
```

```

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 79253: contig of 79253 bp in length
* 79254 79253: gap of unknown length
* 79354 92039: contig of 12686 bp in length
* 92040 92139: gap of unknown length
* 92140 121257: contig of 29118 bp in length.
* Location/Qualifiers
FEATURES             source
    1..121257
     /organism="Medicago truncatula"
     /mol_type="genomic DNA"
     /db_xref="taxon:3880"
     /clone="mth2-9p17"
     /clone_lib="Medicago truncatula BAC library H2"
ORIGIN
Query Match          9.8%; Score 140.2; DB 2; Length 121257;
Best Local Similarity 61.8%; Pred. No. 1.1e-10;
Matches 223; Conservative 0; Mismatches 138; Indels 0; Gaps 0;
QY 33 TGGTCTTGATATCTATGATTTTCAGATTGATTAAGACTTCTATCTATCAGAGACGCC 92
DB 63071 TGAAGCTCAATATCTCTCTCTTATCATGATAAATCTTTCTGTCTACTCTGAGACGCC 63130
QY 93 TGCAGAGGATCCAAATTAGTCTAAATATCTTCAGTCTCGAAACCAACTTCAGGAGCC 152
DB 63131 TTATCTTCTCTGATCATCTTAACCTTTCCGATGTTCTTGACAACTTCCGTCT 63190
QY 153 AAAACCCGTCGTCAACCACTCAGTCTAATATATACAGATGATGACCTTATGACCATAT 212
DB 63191 AACAAAGCACCTCTCCGGACTCAAAACCAACACACGCGTCTTACACTCCGACCATAC 63250
QY 213 AGAGCTCGTAAGTGCCATCTAGATGCGAGATGGAACCTGTTATTTAGCGGACTCA 272
DB 63251 AAGGCTCAACAGTGCCTATCCGATCTAGCATGAAATCTGTTGTTGTAAGTAACTCA 63310
QY 273 ACTAAGCTTAAATATCTCTCAACTCTAGTAAATATACATAGTCTCAATATGTA 332
DB 63311 ATCAACGGTAAACATTCGTCTCAACCAACACCTTCTCTAACAACAGCTCTCAACAA 63370
QY 333 TCCTCTAGTATGATGATCACTTCTCAATTTGACCATCGGTCTGAGGATGGAATGCAGAC 392
DB 63371 TCCTCAACCTGATGATGCTCTCTCCATCTGCCCATCGGTCTGAGGATGGTAAAGCAGAA 63430
QY 393 C 393
DB 63431 C 63431
RESULT 10
AY508219 119270 bp DNA linear HTG 08-MAR-2004
LOCUS Medicago truncatula chromosome 8 clone Mth298J06, *** SEQUENCING IN
DEFINITION PROGRESS ***, 2 ordered pieces.
ACCESSION AY508219
VERSION AY508219.1 GI:42602067
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 119270)
AUTHORS Levy,J., Bres,C., Geurts,R., Chalhouh,B., Kulikova,O., Duc,G.,

```

```

Journet,E.P., Ane,J.M., Lauber,E., Bisseling,T., Denarie,J.,
Rosenberg,C. and Debelle,F.
A putative Ca2+ and calmodulin-dependent protein kinase required
for bacterial and fungal symbioses
Science 303 (5662), 1361-1364 (2004)
14963335
2 (bases 1 to 119270)
Debelle,F., Levy,J., Bres,C., Geurts,R., Chalhouh,B., Kulikova,O.,
Duc,G., Journet,E.-P., Ane,J.-M., Lauber,E., Bisseling,T.,
Denarie,J. and Rosenberg,C.
Direct Submission
Submitted (18-DEC-2003) LIPM, INRA-CNRS, Chemin de Borde Rouge,
BP27, Castanet-Tolosan 31326, France
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 72327: contig of 72327 bp in length
* 72328 72427: gap of unknown length
* 72428 119270: contig of 46843 bp in length.
* Location/Qualifiers
FEATURES             source
    1..119270
     /organism="Medicago truncatula"
     /mol_type="genomic DNA"
     /db_xref="taxon:3880"
     /chromosome="8"
     /clone="Mth298J06"
ORIGIN
Query Match          9.8%; Score 139.8; DB 2; Length 119270;
Best Local Similarity 61.3%; Pred. No. 1.3e-10;
Matches 225; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 27 CCGGGCTGGTCTTGATATCTATGATTTTCAGATTTGATTAAGACTTCTATCTATCAGAA 86
DB 23570 CCAACCTGAATCAATATCTCTCTCTTATCATGATTAATCTTCTGTCTACTCTGA 23629
QY 87 GACGCTGACAGGATCCAAATTAGTCTAAATATTTCTTCTGAGTCTCGGAACCACTCA 146
DB 23630 GACGCTTCATCTTCTCTGAATCATCTTAACCTTTCCGTAGTCTTCTTGACAACTTCC 23689
QY 147 GGACCCAAACCCGTCGCTCACCACTCAGTCTAATATACAGAGTATGACACTTATGA 206
DB 23690 GGTCTTAACAAAGCACTCTCTCCGACTCAAAACCAACATAACGGTGTCTCGACCTCGA 23749
QY 207 CCATATAGAGCTCGTAGGTGCCATCTAGATGCCAGATTGGAACCTGTTATTGTAGCGG 266
DB 23750 CCATACAAAGCTCAACGGTGCCATTCGGATCTAGCATGAAACTCTGTTGTAAGTA 23809
QY 267 AACTCAATAACGGTAAATAATCTCTCAACTACTCTAGTATTAATCAATACATAGCTCAA 326
DB 23810 AACTCAATAACGGTAAACATTCGTCACCACTGACACCTTCTCCAAACCAACAGCTCTC 23869
QY 327 ATCGTATCTCTAGTATGATGATCACTTCTCAATTTGACCATCGGTCTGAGGATGGAAT 386
DB 23870 AACAAATCTCTCAAACTCTGAATAGTCTCTCTGTCTGCCCATCGGTCTGAGGATGGTAA 23929
QY 387 GCAGACC 393
DB 23930 GCAGAAC 23936
RESULT 11
AC149484 96930 bp DNA linear PLN 04-JUN-2004
LOCUS Populus balsamifera subsp. trichocarpa clone Popl-69M19, complete
DEFINITION sequence.
ACCESSION AC149484

```

```
VERSION AC149484.1 GI:48209803
KEYWORDS HTG.
SOURCE Populus balsamifera subsp. trichocarpa (Populus trichocarpa)
ORGANISM Populus balsamifera subsp. trichocarpa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 96930)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 96930)
AUTHORS Stanford Human Genome Center.
CONSRM DOE Joint Genome Institute
TITLE Direct Submission
JOURNAL Submitted (04-JUN-2004) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated total Number of Errors is 0.1.
FEATURES
    source
        1..96930
            /organism="Populus balsamifera subsp. trichocarpa"
            /mol_type="genomic DNA"
            /sub_species="trichocarpa"
            /db_xref="taxon:3694"
            /clone="Popl-69M19"
..
ORIGIN
    Query Match          9.7%;   Score 139;   DB 8;   Length 96930;
    Best Local Similarity 63.1%;   Pred. No. 1.7e-10;
    Matches 214;   Conservative 0;   Mismatches 125;   Indels 0;   Gaps 0;

Qy  52  TTTTCAGATTGTCATAAGACTTCTATCTATCAGAAAGCGCTGCAGAGGATCCCAATTA 111
Db  54757 TTCTTTTATCTGCGTAACACTCTTTTGTCTGCTGAGCTGTTGAAGCTTATTCTTAATTA 54816

Qy  112  GTCTAAATATCTTTCAGTTCGGAACCACTCAGGACCAACCGCTCGTCCACCA 171
Db  54817 CCTCTATCTTCTGAGGTAATCTGAATCAACTCTGTGCCATAGCTCTTCTCACCA 54876

Qy  172  ACTCAGTCTAATATAACAGAGTATGACACTTATGACCATATAGAGCTCTGTAAGGTGCCA 231
Db  54877 CCTCAAACCAACAAACCGGTGATCTACATCTCCAGCATTACAAGCTCATAGGTGCCA 54936

Qy  232  TCTAGATGCCAGATTGGAACACTGTTATTGTAGGCGAACTCACTAACCGTAAATAATCCT 291
Db  54937 TCTCTATGTAGCTGATGATACTGTTGTTGAAGCAAAATCCACAAATGGTAAGAACTTAC 54996

Qy  292  CTCAACTACCTTAGTAATAATCATACATAGCTCCAAATCGTATCTCTAGTATATGAATCA 351
Db  54997 TCCAAACCAACTCCAAAATCCATAACAAAGCCCTTAGCATATCTCCAAGATCTGAATAG 55056

Qy  352  CTTTCTCAAAATTGACCATCGGTCTGAGGATGGAATGCAG 390
Db  55057 TCCTCTCAGATTGACCATCTGCTGAGGGTGAAAAGCTG 55095

RESULT 12
LOCUS AC150760/c 78069 bp DNA linear HTG 10-OCT-2004
DEFINITION Medicago truncatula clone mth2-171i12, WORKING DRAFT SEQUENCE, 15
            unordered pieces.
ACCESSION AC150760
VERSION AC150760.3 GI:54020976
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 78069)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Medicago truncatula BAC Clone mth2-171i12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 78069)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 78069)
AUTHORS Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (10-OCT-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Oct 10, 2004 this sequence version replaced gi:54019504.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 3217: contig of 3217 bp in length
* 3218 3317: gap of unknown length
* 3318 5758: contig of 2441 bp in length
* 5758 5858: gap of unknown length
* 5858 8001: contig of 2143 bp in length
* 8001 8101: gap of unknown length
* 8101 11395: contig of 3294 bp in length
* 11395 11495: gap of unknown length
* 11495 14844: contig of 3349 bp in length
* 14844 14944: gap of unknown length
* 14944 18915: contig of 3971 bp in length
* 18915 19015: gap of unknown length
* 19015 23663: contig of 4648 bp in length
* 23663 23763: gap of unknown length
* 23763 28054: contig of 4291 bp in length
* 28054 28154: gap of unknown length
* 28154 33934: contig of 5780 bp in length
* 33934 34034: gap of unknown length
* 34034 38181: contig of 4147 bp in length
* 38181 38281: gap of unknown length
* 38281 44430: contig of 6149 bp in length
* 44430 44530: gap of unknown length
* 44530 50090: contig of 5560 bp in length
* 50090 50190: gap of unknown length
* 50190 57264: contig of 7074 bp in length
* 57264 57364: gap of unknown length
* 57364 64151: contig of 6787 bp in length
* 64151 64252: gap of unknown length
* 64252 78069: contig of 13818 bp in length.
*
* Location/Qualifiers
* 1..78069
    /organism="Medicago truncatula"
    /mol_type="genomic DNA"
    /db_xref="taxon:3880"
    /clone="mth2-171i12"
    /clone_lib="Medicago truncatula BAC library H2"
ORIGIN
```

Query Match 9.6%; Score 137.6; DB 2; Length 78069;
Best Local Similarity 57.8%; Pred. No. 2.8e-10;
Matches 245; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

Qy 27 CCGGGCTGGCTGATATCATGATTTTCAGATTGCGATGAAGACTTCTATCATACGAA 86
Db 29516 CTTTCTGAAACTCAAGATCTTTTCTAGCTTATCATGATAACTCTTTTGTGCGATTCTGC 29457

Qy 87 GAGCCCTGCGAGGATCCAAATTTAGTCTAAATTTATCTTCAGTCTCGGAAACCAACTCA 146
Db 29456 GAGCCCTTCATTTCTCTGGATCATCGAACTTTCTCAGTAGTCTGCTGAACTCTCT 29397

Qy 147 GGACCCAAACCCGCTGCTACCCCACTCAGTCTTAATATACAGAGATGACACTTATGA 206
Db 29396 GATCCTAAGACCACTCTTTCACTTCACTCAACACGACGACGAGTCTGCACTCCGA 29337

Qy 207 CCATATAGAGCTCGTAAGTGCCATCTAGATGCGAGATTGGAACCTGTTATTGTAGCG 266
Db 29336 CCATATAAGCTCGAAAGGTGCCATTTCCAACTACTAGAAATGATTAATTTATGTATGT 29277

Qy 267 AACTCAACTACGCTGAAATATCTCTCACTTCTCACTTCTAGTAAATCAATAGCTCCAA 326
Db 29276 AACTCGATCAACGGAAGATGCTATCCCAAGTCTCTCTGCTCAAGAACAAATTTCTC 29217

Qy 327 ATCGTATCTCTAGTATATGAATCACCTTCTCAATTTGACCATCGGTCTGAGGATGGAAT 386
Db 29216 AACAAATCTCTAGGACTGAATTTGCTCTTCCAACTGACCATCTGCTCGGATGATAC 29157

Qy 387 GCAGACGGTCCACCGATTACTTAATGGTACTATATAAAATTTATTTTAAAAAA 446
Db 29156 GCCGAATCTCAACTCACTTCCGAACCCAAAGCTCTTGCAAACTCTTCCAAATCTAGAA 29097

Qy 447 TTGA 450
Db 29096 TTA 29093

RESULT 13
AC149480/c
LOCUS AC149480 141416 bp DNA linear PLN 04-JUN-2004
DEFINITION Populus balsamifera subsp. trichocarpa clone Popl-048004, complete sequence.
ACCESSION AC149480
VERSION AC149480.1 GI:48209775
KEYWORDS HTG.
SOURCE Populus balsamifera subsp. trichocarpa (Populus trichocarpa)
ORGANISM Populus balsamifera subsp. trichocarpa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 141416)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
COMMENT 2 (bases 1 to 141416)
AUTHORS Stanford Human Genome Center.
CONSTRM DOE Joint Genome Institute
JOURNAL Direct Submission
Submitted (04-JUN-2004) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
Location/Qualifiers
1. 141416
/organism="Populus balsamifera subsp. trichocarpa"
/mol_type="genomic DNA"
/sub_species="trichocarpa"
/db_xref="taxon:3694"
/clone="Popl-048004"

ORIGIN
Query Match 9.6%; Score 137.4; DB 8; Length 141416;
Best Local Similarity 62.8%; Pred. No. 2.6e-10;
Matches 213; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

Qy 52 TTTTCAGATTGCGATGAAGACTTCTATCTATCAGAAGCGCTGCGAGGATCCCAATTA 111
Db 108532 TTTTATCTCGGTAACTCTTTTGTCTACTCTGAGTTGTTGAAGCTCTCTGTCTAATTA 108473

Qy 112 GTCTAAATTTATCTTCAGTCTCGGAACCAACTCAGGACCCAAACCGTCGCTCACC 171
Db 108472 CCTCTATCTCTCTGAGGTAATCTGAATAAACTCCGGTCCATAGGCTCTTCTCAC 108413

Qy 172 ACTCAGTCTAATAACAGAGATATGACACTTATGACCATATAGAGCTCTCGTAAGTGCCA 231
Db 108412 CCTCAACCAACAAACCGGTGATCTACACTTCGACCATACAAAGCTCTATAAGTGCCA 108353

Qy 232 TCTAGATGCGCAGATTGGAACCTGTTATTGTAGGCGAACTCAACTAAGCGGTAAATAATCCT 291
Db 108352 TCTCTATGCTAGCTGATAACTGTTGTTGAAGCAAAATTCACCTAATGGTAGGAACCTTAC 108293

Qy 292 CTCACCTACCTTAGTAAATTAATCACATAGCTCCAAATCGTATCTCTAGTATATGAATCA 351
Db 108292 TCCAAACCACTCCAAATTCATGACACAGCCCTCAACATATCTCCAAGATCTGAATAG 108233

Qy 352 CCTTCTCAAAATGACCATCGTCTGAGGATGGAATGCAG 390
Db 108232 TCCTCTCAGACTGACCATCTGCTGAGGGTGAAAAGCTG 108194

RESULT 14
AC149299
LOCUS AC149299 143392 bp DNA linear PLN 02-JUN-2004
DEFINITION Populus balsamifera subsp. trichocarpa clone Popl-037B22, complete sequence.
ACCESSION AC149299
VERSION AC149299.1 GI:47900639
KEYWORDS HTG.
SOURCE Populus balsamifera subsp. trichocarpa (Populus trichocarpa)
ORGANISM Populus balsamifera subsp. trichocarpa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 143392)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
COMMENT 2 (bases 1 to 143392)
AUTHORS Stanford Human Genome Center.
CONSTRM DOE Joint Genome Institute
JOURNAL Direct Submission
Submitted (02-JUN-2004) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
Location/Qualifiers
1. 143392
/organism="Populus balsamifera subsp. trichocarpa"
/mol_type="genomic DNA"
/sub_species="trichocarpa"
/db_xref="taxon:3694"
/clone="Popl-037B22"

ORIGIN
Query Match 9.6%; Score 137.2; DB 8; Length 143392;
Best Local Similarity 62.0%; Pred. No. 2.8e-10;
Matches 217; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```
QY 38 CTTGATATCTATGATTTTCAGATTTCGATAAGACTTCTATCTATCAGAAGACGCCCTGCAG 97
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108298 CTCTAAGTCAGCATCTCTCTAATCTGGTAACCTTTTGTCTACTCTGAGCTGTTTGAAG 108357
*
QY 98 AGGATCCCAATATAGTCTAAATATATCTCAGTCTCGGAACCAACTCAGGACCCAAAC 157
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108358 CTTCTTTCTAATTAACCTCTATCTCTCTGAGGTAATCGAATCAACTCTGTGCCATTAG 108417
*
QY 158 CCGTCGCTACCCCAACTCAGCTTAATATACAGAGTATGACACTTATGACCATATAGAGC 217
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108418 CCTCTCTACCAACCTCAACCAACCAACACCGATCTACATCTCGACCATACAAAGC 108477
*
QY 218 CTGCTAGGTGCCATCTAGATGCCAGATTGGAACTGTATTGTAGGCGCAACTCAACTAA 277
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108478 CTCATAAGGTGCCATTTCTATGTAGCTGATCAACTGTTGTGTGAACAAATCCACTAA 108537
*
QY 278 CGGTAAATAAATCTCTCAACTACTCTTAGTATTAATCAATCAGATCTCAATCTGATCTCT 337
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108538 TGGTAGGAATTAATCTCCAACTCACTCCAAATCCAAACCAAGCCCTTAACATATCTTT 108597
*
QY 338 TAGTATATGAATCACCTTCTCAATTCGACCATCGGTCTGAGGATGGAATG 387
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108598 CAAGATCTGAATAGTCTCTCAGACTGACCATCGTCTGAGGTGAAAAG 108647
*
RESULT 15
AC146794/4
LOCUS AC146794 107287 bp DNA linear HTG 14-SEP-2004
DEFINITION Medicago truncatula clone mth2-1019, WORKING DRAFT SEQUENCE, 11
unordered pieces.
ACCESSION AC146794
VERSION AC146794.18 GI:50345165
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Medicago truncatula (Barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 107287)
Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
Medicago truncatula BAC Clone mth2-1019
Unpublished
2 (bases 1 to 107287)
Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
Direct Submission
Submitted (07-OCT-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 107287)
Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
Direct Submission
Submitted (14-SEP-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Jul 16, 2004 this sequence version replaced gi:50057955.
REFERENCE 1
AUTHORS Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Medicago truncatula BAC Clone mth2-1019
JOURNAL Unpublished
REFERENCE 2
AUTHORS Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3
AUTHORS Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Jul 16, 2004 this sequence version replaced gi:50057955.
COMMENT
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 2068: contig of 2068 bp in length
```

```
* 2069 2168: gap of unknown length
* 2169 4879: contig of 2711 bp in length
* 4880 4979: gap of unknown length
* 4980 7400: contig of 2421 bp in length
* 7401 7500: gap of unknown length
* 7501 13700: contig of 6200 bp in length
* 13701 13800: gap of unknown length
* 13801 16982: contig of 3182 bp in length
* 16983 17082: gap of unknown length
* 17083 22047: contig of 4965 bp in length
* 22048 22147: gap of unknown length
* 22148 27289: contig of 5142 bp in length
* 27290 27390: gap of unknown length
* 27391 41538: contig of 14149 bp in length
* 41539 41638: gap of unknown length
* 41639 54952: contig of 13314 bp in length
* 54953 55052: gap of unknown length
* 55053 81486: contig of 26434 bp in length
* 81487 81587: gap of unknown length
* 81588 107287: contig of 25701 bp in length.
FEATURES
Location/Qualifiers
1..107287
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-1019"
/clone_lib="Medicago truncatula BAC library H2"
ORIGIN
Query Match 9.5%; Score 136.4; DB 2; Length 107287;
Best Local Similarity 61.0%; Pred. No. 3.8e-10;
Matches 221; Conservative 0; Mismatches 141; Indels 0; Gaps 0;
QY 45 TCTATGATTTTCAGATTTCGATAAGACTTCTATCTATCAGAAGACGCTTCAGAGGATCC 104
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84701 TCTTTCTAATGCTTGTGCGTGAATCACTTATATGTCGACTCTGCGACGCTTTCATCTTTCT 84642
*
QY 105 CAAATTAGTCTAAATATTTCTTCAGTCTCGGAACCAACTCAGGACCCAAACCCGTCGC 164
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84641 CTTAATAGCTTAACTTCTCTGTAGTCTGGTGAACAATATTCGGTCCCAACTACACTC 84582
*
QY 165 TCACCCCACTCAGTCTTAATATACAGATGATGACATTCATGACCATATAGAGCTCGTAA 224
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84581 TCACCCGACTCAAAACCAACATAACAGAGTTCTCGACCTCGGACCATACAAGCTTCGAAT 84522
*
QY 225 GGTGCCATCTAGATGCCAGATTGGAAACTGTTATTGTAGGCGAACTCAACTAAGCTTAA 284
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84521 GGTGCCATCTCTATACAGATGATGATGTTGTTGTACGTGAACCTATCAACGGTAGA 84462
*
QY 285 AAATCTCTCAACTACTCTTAGTAAATAACATACATAGCTCCAAATCGTATCTCTAGTATA 344
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84461 TGACTATCCCAAGTCCCACTTGTATCAAGTACACAACTCTCAACAATCTCTTAACGAC 84402
*
QY 345 TGAATCCTTCTCAATTCAGCATCGCTGAGGATGGAATGCAGACCGGTGCACCGA 404
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84401 TGAATCGTCTCTCTGACTGACCATCTGTTGGGATGATAGCCGAATCAACCTCAAC 84342
*
QY 405 TT 406
|||
Db 84341 TT 84340
|||
Search completed: August 13, 2005, 11:00:00
Job time : 10343 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2005, 02:09:43 ; Search time 7384 Seconds
(without alignments)

7387.066 Million cell updates/sec

Title: US-10-089-543-2

Perfect score: 1433

Sequence: 1 actataggcagcggtgttc.....tgagagaatccttcacatc 1433

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_hc.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_gse1.*

9: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	224.4	15.7	581	9	CL864797 TM1-GSS00
2	183.8	12.8	768	9	CL251651 ZMMBBb060
3	175.4	12.2	758	9	CL404830 ZMMBBb039
4	172.8	12.1	1006	9	CL263790 ZMMBBb062
5	172	12.0	447	9	CL864019 TM1-GSS00
6	164.6	11.5	973	9	CL249359 ZMMBBb059
7	162.4	11.3	762	9	CL863878 TM1-GSS00
8	157.6	11.0	643	9	CL864239 TM1-GSS00
9	155.8	10.9	355	9	CL202255 ZMMBBb056
10	151.6	10.6	1057	9	CL285168 ZMMBBb062
11	147.4	10.3	957	9	CL291015 ZMMBBb063
12	146.4	10.2	739	9	CL864139 TM1-GSS00
13	143.2	10.0	569	8	BH021651 GH_MBB000
14	142	9.9	686	8	BH022375 GH_MBB000
15	139.8	9.8	725	9	CC746178 ZMMBBb012
16	139.8	9.8	781	8	BH022260 GH_MBB000
17	139.6	9.7	750	9	CC688457 OGWC37TH
18	139.6	9.7	1015	9	CL010422 ZMMBBb055
19	137.4	9.6	509	4	B1122809 I013P31P
20	137.4	9.6	799	8	BZ702320 PUCD182TD
21	137.2	9.6	660	8	BH021854 GH_MBB000
22	137	9.6	694	8	BZ420292 if66d04.b
23	136.8	9.5	697	8	BH985303 oeil6d03
24	136.8	9.5	843	9	CG006002 ZUACL81TV

25	136.4	9.5	679	9	CL243740
26	136.4	9.5	735	9	CC752373
27	136.4	9.5	786	9	CC757350
28	136.4	9.5	893	9	CC979403 ZUAGY07TH
29	136.4	9.5	931	9	CC978577 ZUAGY07TH
30	136.4	9.5	957	9	CL248255
31	135.8	9.5	750	9	CC783903
32	135.8	9.5	792	9	CC986980
33	135.8	9.5	818	9	CC815980
34	135.8	9.5	848	9	CG245637
35	135.8	9.5	857	9	CG909847
36	135.2	9.4	651	8	BZ488595
37	135.2	9.4	653	9	CC962470
38	135.2	9.4	708	9	CC760866
39	135.2	9.4	797	8	BZ457826
40	135.2	9.4	805	8	BZ453701
41	134.8	9.4	629	8	BZ719317
42	134.8	9.4	671	9	CG867526
43	134.8	9.4	787	9	CC603679
44	134.6	9.4	951	9	CG300841
45	134.2	9.4	705	8	CC146564

ALIGNMENTS

RESULT 1
LOCUS CL864797
DEFINITION TM1-GSS000073r BAC and BIBAC libraries from Upland cotton genetic standard TM-1 Gossypium hirsutum genomic clone GH-TM1-CBV149b20 5', genomic survey sequence.

ACCESSION CL864797
VERSION CL864797.1
KEYWORDS GI:51321527
SOURCE GSS.
ORGANISM Gossypium hirsutum (upland cotton)

REFERENCE 1 (bases 1 to 581)
AUTHORS Xu,Z., Kohel,R.J., Zhang,H.B., Dong,J., Covalada,L., Lee,M., Koo,P. and Yu,J.Z.
TITLE Genome-Wide Synteny between Arabidopsis and Cotton
JOURNAL Unpublished (2004)
COMMENT Other GSSs: TM1-GSS000073f

Contact: John Z. Yu
Cotton Molecular Genetics Laboratory
USDA-ARS Crop Germplasm Research Unit
2765 F&B Road, Building 11, College Station, TX 77845. USA
Tel: 979 -260-9237
Fax: 979-260-9333
Email: zyu@qutun.tamu.edu
for more detail, please see
http://algonon.tamu.edu/htdocs-cotton/cottondb.html
Seg primer: sp030
Class: BAC ends
High quality sequence stop: 581.

FEATURES
source location/Qualifiers
1..581
/organism="Gossypium hirsutum"
/mol_type="genomic DNA"
/cultivar="TM-1"
/db_xref="taxon:3635"
/clone="GH-TM1-CBV149b20"
/tissue_type="young leaves"
/lab_host="DH10B"
/clone_lib="BAC and BIBAC libraries from Upland cotton genetic standard TM-1"
/note="vector: pCLD04541; pBeloBAC11; For more details on library construction, ordering clones and sequence analysis see
http://algonon.tamu.edu/htdocs-cotton/cottondb.html"

CL243740 ZMMBBb039
CC752373 ZMMBBb013
CC757350 ZMMBBb014
CC979403 ZUAGY07TH
CC978577 ZUAGY07TH
CL248255 ZMMBBb059
CC783903 ZMMBBb014
CC986980 ZUAD810TH
CC815980 ZMMBBb052
CG245637 OGXB12TV
CG909847 ZMMBBb034
BZ488595 BONY26TF
CC962470 BOIGF89TR
CC760866 ZMMBBb015
BZ457826 BONLD73TF
BZ453701 BONFSS5TF
BZ719317 PUCFH65TD
CG867526 ZMMBBb035
CC603679 OGWB548TV
CG300841 OGOC059TV
CC146564 ZMMBBb000

ORIGIN

Query Match 15.7%; Score 224.4; DB 9; Length 581;
Best Local Similarity 77.1%; Pred. No. 3.1e-27;
Matches 273; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 36 TGCTTGATATCATGATTTTCAGATTTGCATAAGACTTCTATCTATCATCAGAAAGCGCTGC 95
Db 195 TACTCAATATCATGCTTTTTCAGATTTGCATAAGACTTCTATCTATCATCAGAAAGCGCTGC 254

Qy 96 AGAGGATCCCAATAGTCTAAATATCTTCAATCTCGGAAACCAACTCAGGACCCAAA 155
Db 255 AGACGATCTCAACACAGTTTAACTTGCCTTCAGTCTCAAAAACCAACTCAGGACCCAAA 314

Qy 156 ACCGTCGCTCACCCTAGTCTAATAAACAAGATGACACTTTATGACCAATATAGA 215
Db 315 ACTTGATGCTTACCCTACTGCTCTAATAGTAGTGAATAGACACTTACCAACATATAAA 374

Qy 216 GCCTCGTAAGGTGCATCTAGATGCCAGATTGGAACTGTATTGTAGGCGAACTCAACT 275
Db 375 GCCTCGTAAGCGCATATGATGATGCTAGACTAGAACTGTATTGTAGGCGAACTCAGCT 434

Qy 276 AACCGTAAAAAATCCTCTCAACTACCTTAGTAATAATCATAGTCTCCAAATCGTATCC 335
Db 435 AGTGGCAGATAATCTCCCACTACCTCAGAAATCAATCACAAATTTTGAAGTATATCC 494

Qy 336 TCTAGTATGAATCACTCTCAAAATGACCATCGGTCTGAGGATGGAATGCA 389
Db 495 TTTGGTATCTGAATCACCTCTTAGATTGACACCGGTCTGAGGATGGAACGCA 548

RESULT 2

CL251651/CL251651/c

LOCUS ZM5B0604M16f ZM5B5b (HindIII) Zea mays genomic clone

DEFINITION ZM5B5b0604M16 5', genomic survey sequence.

ACCESSION CL251651

VERSION CL251651.1 GI:41354780

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 768)

AUTHORS Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C., Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.

TITLE Sequencing of the maize genome at PGR (2003c)

JOURNAL Unpublished (2003)

COMMENT Contact: Bharti,A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 430.
Location/Qualifiers
1. .768
/organism="Zea mays"
/mol_type="genomic DNA"
/cultiivar="B73"
/db_xref="taxon:4577"
/clone="ZM5B5b0604M16"
/lab_host="E. coli DH10B"
/clone_lib="ZM5B5b (HindIII)"
/note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 12.8%; Score 183.8; DB 9; Length 768;

Best Local Similarity 74.8%; Pred. No. 1.5e-20;
Matches 270; Conservative 0; Mismatches 87; Indels 4; Gaps 3;

Qy 39 TTGATATCTATGATTTTTCAGATTTGCATAAGACTTCTAT-CTATCAGAAAGCGCTGCAG 97
Db 424 TAGATATCCCAATTTCTCAGATCTGAATAAGATATATATATATCAGAGCTGCTTAAG 365

Qy 98 AGGATCCCAAAATPAGTCTAAAAATTTCTTCAGTCTCGGAAACCAACTCAGGAC-CCAAAA 156
Db 364 AGGATCCCAAAATCATTTCTAACTTTTATCTCGTTCGGTTTCAGAAACCAACTCAGAGACAAAA 305

Qy 157 CCGTCGCTCACCCTAGTCTATATATAACAGAGTATGACACTTTATGACCAATATAGAG 216
Db 304 TCTGTGCTCGCTCGCCCACTTAGTTTCAACAATAAGTACGACACTTACGACCATACAGAG 245

Qy 217 CCGTGAAGGTGCATCTAGATGCCAGATTGGAACTGTATTGTAGGCGAACTCAACTA 276
Db 244 CCGTGAAGGTGCTATTGTAATCTAGACTGGAACCTATTATTATAAGTGAACCTCGACTA 185

Qy 277 ACAGTAAAAAATCCTCTCAACTACCTTAGTAAATCAATCAATGCTCCAAATCGTATCCT 336
Db 184 ACAGTAAAAAATCCTCTCAAGACTACCTCGAAAGTCAATCACAACTCCAAAAACATATCTT 125

Qy 337 CTAGTATATGAATCACTCTTCT--CAATTTGACCATCGGTCTGAGGATGGAATCGACACCG 394
Db 124 CCAGTATCTAAATCACCTCTCCCTAAGAAAGATCTATCTAAGGATGGAATGTAGTACT 65

Qy 395 G 395
Db 64 G 64

RESULT 3

CL404830/CL404830/c

LOCUS ZM5B5b0397006f ZM5B5b (HindIII) Zea mays genomic clone

DEFINITION ZM5B5b0397006 5', genomic survey sequence.

ACCESSION CL404830

VERSION CL404830.1 GI:45216049

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 758)

AUTHORS Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C., Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.

TITLE Sequencing of the maize genome at PGR (2003c)

JOURNAL Unpublished (2003)

COMMENT Contact: Bharti,A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 89.
Location/Qualifiers
1. .758
/organism="Zea mays"
/mol_type="genomic DNA"
/cultiivar="B73"
/db_xref="taxon:4577"
/clone="ZM5B5b0397006"
/lab_host="E. coli DH10B"
/clone_lib="ZM5B5b (HindIII)"
/note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 12.2%; Score 175.4; DB 9; Length 758;
 Best Local Similarity 75.1%; Pred. No. 3.6e-19;
 Matches 259; Conservative 0; Mismatches 81; Indels 5; Gaps 3;

QY 52 TTTTCAGATTGCTAAGACTTCTATCTATCAGAGAGCGCTGAGAGGATCCAAATTA 111
 |||||
 DB 750 TATTCAAATTTGCATATGA-TTTTGTCTATCAGAAGCGCTTTTCAAACGATTCGAATCA 692
 |||||

QY 112 GTCTAAATATCTTTCAGTCTCGGAAACCAACTCAGGACCCAAACCGCTCGCTCACCCA 171
 |||||
 DB 691 ATCTAACCTCATCTCTATCTCAGAGACCAACTCTGGACCCAGAACAGCTGCTCGCCCA 632
 |||||

QY 172 ACTCAGT-CTAATAAACAAGATGACACTTATGACCATATAGAGCCTCGTAAGTGCC 230
 |||||
 DB 631 ACTTAGTCCCAACGTAAGGAGTGGACACTTACAACCGTACAGTGCCTCGTAAGTGCC 572
 |||||

QY 231 ATCTAGATGCCAGATTGGAAGTCTTATTTAGTGGGAACTCAACTACGGTAAAAATCC 290
 |||||
 DB 571 ATCTGGATGCTAGACTGGAAGCTATTATTTCTAGGCAAACTCTGTAA---TGCAAGTAC 515
 |||||

QY 291 TCTCAACTACCTTAGTAATAATCACATAGTCTCCAAATCGTATCTCTAGTATATGAATC 350
 |||||
 DB 514 TCCCACTACATCAAGAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 455
 |||||

QY 351 ACCTTCTCAAATTCAGCATCGTCTGAGGATGGAATGCAGACCGG 395
 |||||
 DB 454 ATCTCTCTAACTGAGCGTCTGTTTGGAGTGAATAACAGTACTG 410
 |||||

RESULT 4
 CL263790 1006 bp DNA linear GSS 02-FEB-2004
 LOCUS ZMWB0623P14f ZMWB0623P14 5', genomic survey sequence.
 DEFINITION ZMWB0623P14 5', genomic survey sequence.

ACCESSION
 VERSION CL263790.1 GI:41627929
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 1006)
 Bharti,A.K., Young,S., Kaychok,S., Keizer,G., Bronzino,A.C., Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
 Sequencing of the maize genome at PGR (2003c)
 Unpublished (2003)
 CONTACT: Bharti,A.K.
 Dr.Joschim Messing's lab
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 Tel: 732 445 3801
 Fax: 732 445 5735
 Email: bharti@waksman.rutgers.edu

Seq primer: T7
 Class: BAC ends
 High quality sequence start: 51.
 Location/Qualifiers
 1..1006
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMWB0623P14"
 /lab_host="E. coli DH10B"
 /clone_lib="ZMWB0623P14"
 /note="vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

FEATURES
 source
 1..1006
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMWB0623P14"
 /lab_host="E. coli DH10B"
 /clone_lib="ZMWB0623P14"
 /note="vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

ORIGIN
 Query Match 12.1%; Score 172.8; DB 9; Length 1006;
 Best Local Similarity 69.7%; Pred. No. 9.2e-19;
 Matches 248; Conservative 0; Mismatches 107; Indels 1; Gaps 1;

QY 36 TGCTTGATATCTATGATTTTTCAGATTGTCATAAGACTTCTATCTATCAGAGAGCGCTGC 95
 |||||
 DB 96 TACTCAATCTCTTGGCATTTTAAGATCGCGTAAGACTGCTACCTATCAGATGCTTCTCTC 155
 |||||

QY 96 AGAGATCCCAATTTAGTCTAAAATTTATCTTTCAGTCTCGGAAACCAACTCAGGACCCAAA 155
 |||||
 DB 156 AGTTGGTCTCGAATCAGTTTACCTTTCTCTTCGGTATCAGAAAATAACTCTGGCCCTAGA 215
 |||||

QY 156 ACCGTCGCTCACCCAACTCAGTCTAATAATAACAGAGTATGACACTTATGACCATATAGA 215
 |||||
 DB 216 ACTCGTCTCGGCCAACTCAGTCCAAACAGTAGGATGACACATACACCATATATAC 275
 |||||

QY 216 GCCTCGTAAAGTGCCATCTAGATGCCAGATTGGAACCTGTTATTTAGTGGGAACTCAACT 275
 |||||
 DB 276 GCCTCGTACGCTGCTATTTTGAATCTGCTGCTGGTAGCTATTTATTAAGCGACTCTACT 335
 |||||

QY 276 AACGGTAAAAATCTCTCTCAACTACCTTAGTATTAATATCAC-ATAGTCCAAATCGTATC 334
 |||||
 DB 336 AGCGGCAAAATAGTCTTTCCAGGTGCTCGGAAATCCAAATCTACTTATTTTCTTAACATATC 395
 |||||

QY 335 CTCTAGTATATGAATCACCTTCTCAATTCACCATCGTCTGAGGATGGAATGCAG 390
 |||||
 DB 396 CTCATATTTTGAATCTCTCTTCTGATGACCATCTATCTGGGATGAACGCGAG 451
 |||||

RESULT 5
 CL864019/c 447 bp DNA linear GSS 19-AUG-2004
 LOCUS CL864019/c BAC and BIBAC libraries from Upland cotton genetic standard TM-1 Gossypium hirsutum genomic clone GH-TM1-CBV047P09 5', genomic survey sequence.

ACCESSION
 VERSION CL864019.1 GI:51320749
 KEYWORDS GSS.
 SOURCE Gossypium hirsutum (upland cotton)
 ORGANISM Gossypium hirsutum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 447)
 Xu,Z., Koneh,R.J., Zhang,H.B., Dong,J., Covalada,L., Lee,M., Koo,P. and Yu,J.Z.
 Genome-Wide Synteny between Arabidopsis and Cotton
 Unpublished (2004)
 Other GSSs: TM1-GSS000206r
 CONTACT: John Z. Yu
 Cotton Molecular Genetics Laboratory
 USDA-ARS Crop Germplasm Research Unit
 2765 F&B Road, Building 11, College Station, TX 77845. USA
 Tel: 979 -260-9237
 Fax: 979-260-9333
 Email: zyu@utmsi.tamu.edu
 for more detail, please see
 http://algodon.tamu.edu/htdocs-cotton/cottondb.html

Seq primer: sp010
 Class: BAC ends
 High quality sequence stop: 447.
 Location/Qualifiers
 1..447
 /organism="Gossypium hirsutum"
 /mol_type="genomic DNA"
 /cultivar="TM-1"
 /db_xref="taxon:3635"
 /clone="GH-TM1-CBV047P09"
 /tissue_type="young leaves"
 /lab_host="DH10B"
 /clone_lib="BAC and BIBAC libraries from Upland cotton genetic standard TM-1"
 /note="vector: pCLD04541; pBeloBAC11; For more details on library construction, ordering clones and sequence analysis see
 http://algodon.tamu.edu/htdocs-cotton/cottondb.html"

FEATURES
 source
 1..447
 /organism="Gossypium hirsutum"
 /mol_type="genomic DNA"
 /cultivar="TM-1"
 /db_xref="taxon:3635"
 /clone="GH-TM1-CBV047P09"
 /tissue_type="young leaves"
 /lab_host="DH10B"
 /clone_lib="BAC and BIBAC libraries from Upland cotton genetic standard TM-1"
 /note="vector: pCLD04541; pBeloBAC11; For more details on library construction, ordering clones and sequence analysis see
 http://algodon.tamu.edu/htdocs-cotton/cottondb.html"

ORIGIN
 Query Match 12.1%; Score 172.8; DB 9; Length 1006;
 Best Local Similarity 69.7%; Pred. No. 9.2e-19;
 Matches 248; Conservative 0; Mismatches 107; Indels 1; Gaps 1;

ORIGIN	Query Match	12.0%;	Score 172;	DB 9;	Length 447;
	Best Local Similarity	67.7%;	Pred. No. 1.4e-18;		
	Matches 241;	Conservative 0;	Mismatches 115;	Indels 0;	Gaps 0;
QY	35	GTGCTTGATATCTATGATTTTCAGATTTTGCATATAGACTTCTATCTATCAGAAGCGCTG	94		
DB	417	GTATTCAATCTCTTTACGCTTCAGATCTGCATACGACTTTTGGCTTATCAGACGCTTCTCT	358		
QY	95	CAGAGATCCCAAAATTAGTCTAAAAATTATCTTCAGTCTCGGAAACCAACTCAGGACCCAA	154		
DB	357	TAATCGATCTTAATTAACCTGACCTTATCTCTAGTATCTGCCACCACTCAAGTCCAA	298		
QY	155	AACCCGTGCTCACCACCACTCAGTCTAATATAACAGAGATGACATTTATGACCATATAG	214		
DB	297	AACTTGTGCTCCCCCAGTTCAGTCCAAACAATAGGTGTTTGACACCTTCGCCCCACACAG	238		
QY	215	AGCCTCGTAAGTGGCCATCTAGATGCCAGATTGGAAACGTGTTATGTAGGCGGAATCAAC	274		
DB	237	TGCTTTCATAAGTGGCATCCGAAATACTCGACTGATAACTTGTGTATGATAAACTCTTC	178		
QY	275	TAAACGTAATAAAATCTCTCAACTACCTTAGTAAATAAAATCACATAGCTTCCAAATCGTATC	334		
DB	177	CAAGGCAAGTAGTCTCTCCCACTACCTCGGAAATCAATAACACATCCCTCAGCATGTC	118		
QY	335	CTCTAGTATATGAATCACTCTCATAATTGACCATCGGTCTGAGGATGGATGCA	390		
DB	117	TTCCAGAACTCGAATAACTCTTTCCGACTGACCATCAGTTTGGGATGCGAAAGCG	62		

```

RESULT 6 .
.....
CL249359/c
LOCUS           973 bp      DNA      linear      GSS 22-JAN-2004
DEFINITION      ZMMBBb0595C21f ZMMBBb (HindIII) Zea mays genomic clone
                ZMMBBb0595C21 5', genomic survey sequence.
ACCESSION       CL249359
VERSION         CL249359.1  GI:41105913
KEYWORDS        GSS.
SOURCE          Zea mays
ORGANISM        Zea mays
REFERENCE       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
AUTHORS         Bharti A.K., Young S.G., Kavchok S., Keizer G., Bronzino A.C.,
                Zohoverz V., Fuku G., Yu Y., Wing R. and Messing J.
TITLE           Sequencing of the maize genome at PGIR (2003C)
JOURNAL         Unpublished (2003)
COMMENT         Contact: Bharti, A.K.
                Dr. Joachim Messing's lab
                The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
                University
                190 Frelinghuysen Road, Piscataway, NJ 08854, USA
                Tel: 732 445 3801
                Fax: 732 445 5735
                Email: bharti@waksman.rutgers.edu
                Seq primer: T7
                Class: BAC ends
                High quality sequence start: 50.
FEATURES        1..973
                source          /organism="Zea mays"
                                /mol_type="genomic DNA"
                                /cultivar="B73"
                                /db_xref="taxon:4577"
                                /clone="ZMMBBb0595C21"
                                /lab_host="E. coli DH10B"
                                /clone_lib="ZMMBBb (HindIII)"
                                /note="Vector: pCGU1; Site_1: HindIII; Site_2: HindIII"
ORIGIN
Query Match    11.5%; Score 164.6; DB 9; Length 973;

```


Tel: 732 445 3801
 Fax: 732 445 5735
 Email: bharti@waksman.rutgers.edu
 Seq primer: SP6
 Class: BAC ends
 High quality sequence start: 180.
 Location/Qualifiers
 1. .355
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBB063L12"
 /lab_host="E. coli DH10B"
 /clone_lib="ZMMBBb (HindIII)"
 /note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

FEATURES

source

ORIGIN

Query Match 10.9%; Score 155.8; DB 9; Length 355;
 Best Local Similarity 81.7%; Pred. No. 6.7e-16;
 Matches 192; Conservative 0; Mismatches 42; Indels 1; Gaps 1;
 Qy 157 CCGTCCTCACCACCTCAGTCTAATATATACAGAGTATGACACTTATGACCATATAGAG 216
 Db 355 CCGGACGCTCACCACCTTATGTCACACATAGCAGAGTTCTACACTTACGATCATACAAG 296
 Qy 217 CCTCGTAAGTGCCATCTAGATCCAGATTGGAACTGTTATTGTAGGCGAACTCAACTA 276
 *Db 295 CCTCGTAAGTGCCATCTGGATGCTAGCTAGCAAGAACTGTTATTGTAGGCGAACTCAACTA 236
 Qy 277 ACGGTAAAAATCCTCTCAACTACCTTAGTAATAAATCACATAGCTCCAAATCGTATCCT 336
 Db 235 GCGGTAGTAACTCTCCCACTACCTCTGAAATCTATTACACAACCTGGAAGCATATCCT 176
 Qy 337 CTAGTATATGAATCACTCTTCAAAATGACCATCGGT-CTGAGATGGAATGCAG 390
 Db 175 CTATAATTGAATCACCCTCTCAGACTGACCCTCGGTACTAAGGATGGAACGTAG 121

RESULT 10

CL285168

LOCUS

DEFINITION

ZMMBB0628K19r ZMMBBb (HindIII) Zea mays genomic clone

ZMMBB0628K19 3', genomic survey sequence.

CL285168

GSS.

KEYWORDS

SOURCE

ORGANISM

Zea mays

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Bharti, A.K.

Dr. Joachim Messing's lab

The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
 University

190 Frelinghuysen Road, Piscataway, NJ 08854, USA

Tel: 732 445 3801

Fax: 732 445 5735

Email: bharti@waksman.rutgers.edu

Seq primer: SP6

Class: BAC ends

High quality sequence start: 405.

Location/Qualifiers

1. .1057

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"
 /clone="ZMMBBb0628K19"
 /lab_host="E. coli DH10B"
 /clone_lib="ZMMBBb (HindIII)"
 /note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 10.6%; Score 151.6; DB 9; Length 1057;
 Best Local Similarity 68.3%; Pred. No. 2.9e-15;
 Matches 239; Conservative 0; Mismatches 109; Indels 2; Gaps 2;
 Qy 43 TATCTATGATTTTCAGATTGCGATAGACTTCTATC-TATCAGAAGACCGCTGCAGAGA 101
 Db 78 TCTCTTTTATGTTTAAATCTGTATACGACCTGACTCAGGGCAGATCTCTTTTCAATCGG 137
 Qy 102 TCCCAATTTAGTCTAAATTTCTTTCAGTCTCGGAAACCAACTCAGGACCCAAACCCGT 161
 Db 138 TCTCGAATCAGTTTACCTTATCTTCGGTATCAGAACTAACTCTGGGCTAGAACTCGC 197
 Qy 162 CGCTCACCAACTCAGTCTAATATATACAGAGTATGACATTATGACCATATAGAGCTCG 221
 Db 198 CACTCGCTTAGCTCAGTCCCAACAATAGGAGTATGACACTATGACCATATAATATGCTCG 257
 Qy 222 TAAGTGCCATCTAGATGCCAGATTGGAACCTGTTATTCTAGCGCAACTCACTACGCT 281
 Db 258 TAGGATGCCATTTCAATACTGGAATCTGTTATTTATATATAGCAAACTTCTGCTAGCGC 317
 Qy 282 AAAAAATCTCTCAACTACCTTAGTAATAAATCACATAGCTCCAAATCGTATCTCTAGT 341
 Db 318 AATATATCTTCCAAATGCTTGGAAATCCATTACACAACCTCAACATATCTCCAGT 377
 Qy 342 ATATGAATCACTTCT-CAAAATGACCATCGGTCTGAGATGGAATGCAG 390
 Db 378 ACTTGAATCACTTTTTCCGATTGTTCTATCTGCTGGGATGACAGCGAG 427

RESULT 11

CL291015/c

LOCUS

DEFINITION

ZMMBB0636D02r ZMMBBb (HindIII) Zea mays genomic clone

ZMMBB0636D02 3', genomic survey sequence.

CL291015

GSS.

KEYWORDS

SOURCE

ORGANISM

Zea mays

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Bharti, A.K.

Dr. Joachim Messing's lab

The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
 University

190 Frelinghuysen Road, Piscataway, NJ 08854, USA

Tel: 732 445 3801

Fax: 732 445 5735

Email: bharti@waksman.rutgers.edu

Seq primer: SP6

Class: BAC ends

High quality sequence start: 399.

Location/Qualifiers

1. .957

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="ZMMBBb0636D02"

/lab_host="E. coli DH10B"

Class: BAC ends.
Location/Qualifiers
1. .725
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBBb0122010"
/lab_host="DH10B"
/clone_lib="ZMMBBb"
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2:
HindIII; Zea mays L. ssp. mays"

ORIGIN
Query Match 9.8%; Score 139.8; DB 9; Length 725;
Best Local Similarity 68.1%; Pred. No. 2.7e-13;
Matches 209; Conservative 0; Mismatches 97; Indels 1; Gaps 1;
Qy 80 ATCAGAAGACGCTGCAGAGATCCCAAATTAGTCTAAATTTATCTTCAGTCTCGAAAC 139
Db 308 ATCCGAGCTGCTTCAGAGTATAACTTATCATTTTAACCTTTTCTTCAGTTTGTGCTAT 249
Qy 140 CAATCTCAGGACCCAAACCCGCTCGCTACCCAACTCAGTCTTAATATAACAGAGATGACA 199
Db 248 CAAATCAACCCGGAATACCTTTTCTCCCTGAGCTCAGTCTATTATAACAAAGTTGACA 189
Qy 200 CTTATGACCATATAGAGCTCGTAAGGTGCCATCTAGATGCCAGATTGGAACCTGTTATT 259
Db 188 CTTCTGACCATACATGCTTCATACATTTGTCATTTTAAGCTCGATTGATAACTGTTTTT 129
Qy 260 GTAGCGCAACTCAACTAACGGTAAAAATCCTCTCACTACCTTAGTAAATAAATCACATA 319
Db 128 ATATGCGAATTCAACTAATGTAGGAATCTCTCCCACTACCTTAAATTTCTAAACACA 69
Qy 320 GCTCCAAATCGTATCCTCTAGTATATGAATCA-CCTTCTCAAAATGACCATCGGCTCGAG 378
Db 68 ATGTGGAAGCATATCTTCAAGTATTGGAATCATAGGATTGAGATTGACCATCTGCTGGG 9
Qy 379 GATGGA 385
Db 8 GATGAA 2

Search completed: August 13, 2005, 13:05:14
Job time : 7392 secs

THIS PAGE BLANK (U)